Vol. 35, No. 3 May, 1928

PSYCHOLOGICAL REVIEW PUBLICATIONS

Psychological Review

EDITED BY

HOWARD C. WARREN, PRINCETON UNIVERSITY
MADISON BENTLEY, UNIVERSITY OF ILLINOIS (*J. of Exper. Psychol.*)
S. W. FERNBERGER, UNIV. OF PENNSYLVANIA (*Bulletin*)
W. S. HUNTER, CLARK UNIVERSITY (*Index*), AND
RAYMOND DODGE, YALE UNIVERSITY (*Monographs*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY, *Business Editor*

ADVISORY EDITORS

R. P. ANGIER, YALE UNIVERSITY; MARY W. CALKINS, WELLESLEY COLLEGE; JOSEPH JASTROW, UNIVERSITY OF WISCONSIN; C. H. JUDD, UNIVERSITY OF CHICAGO; ADOLF MEYER, JOHNS HOPKINS UNIVERSITY; W. B. PILLSBURY, UNIV. OF MICHIGAN; C. E. SEASHORE, UNIVERSITY OF IOWA; G. M. STRATTON, UNIV. OF CALIFORNIA; MARGARET F. WASHBURN, VASSAR COLLEGE; JOHN B. WATSON, NEW YORK; R. S. WOODWORTH, COLUMBIA UNIVERSITY.

CONTENTS

The Absolute Zero in Intelligence Measurement: L. L. THURSTONE, 175.

Intelligence as the Capacity for Variability or Versatility of Response: A. S. EDWARDS, 198.

The Theory of Factors, I: STUART C. DODD, 211.

The Inference of Mind: DONALD KEITH ADAMS, 235.

A Further Experimental Study of the Development of Behavior: LEONARD CARMICHAEL, 253.

PUBLISHED BI-MONTHLY
FOR THE AMERICAN PSYCHOLOGICAL ASSOCIATION
BY THE PSYCHOLOGICAL REVIEW COMPANY
PRINCE AND LEMON STS., LANCASTER, PA.
AND PRINCETON, N. J.

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879

Psychological Review Publications

OF THE

American Psychological Association

EDITED BY

HOWARD C. WARREN, PRINCETON UNIVERSITY (*Review*)
RAYMOND DODGE, YALE UNIVERSITY (*Monographs*)
SAMUEL W. FERNBERGER, UNIV. OF PENNSYLVANIA (*Bulletin*)
W. S. HUNTER, CLARK UNIVERSITY (*Index*)
MADISON BENTLEY, UNIVERSITY OF ILLINOIS (*J. of Exp. Psych.*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY, *Business Editor*

WITH THE COOPERATION OF
MANY DISTINGUISHED PSYCHOLOGISTS

PSYCHOLOGICAL REVIEW

containing original contributions only, appears bimonthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 480 pages.

PSYCHOLOGICAL BULLETIN

containing critical reviews of books and articles, psychological news and notes, university notices, and announcements, appears monthly, the annual volume comprising about 720 pages. Special issues of the BULLETIN consist of general reviews of recent work in some department of psychology.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

containing original contributions of an experimental character, appears bimonthly, February, April, June, August, October, and December, the six numbers comprising a volume of about 480 pages.

PSYCHOLOGICAL INDEX

is a compendious bibliography of books, monographs, and articles upon psychological and cognate topics that have appeared during the year. The INDEX is issued annually in May, and may be subscribed for in connection with the periodicals above, or purchased separately.

PSYCHOLOGICAL MONOGRAPHS

consists of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The MONOGRAPHS appear at irregular intervals and are gathered into volumes of about 500 pages.

ANNUAL SUBSCRIPTION RATES

Review: $5.00 (Foreign, $5.25).
Journal: $5.00 (Foreign, $5.25).
Bulletin: $5.50 (Foreign, $5.75).
Review and Bulletin: $9.50 (Foreign, $10.00).
Review and Journal: $9.00 (Foreign, $ 9.50).
Journal and Bulletin: $9.50 (Foreign, $10.00).

Any one of above with Index: $1.50 additional.

Review, Bulletin, and Journal: $14.00 (Foreign, $14.75).
Review, Bulletin, Journal, and Index: $15.00 (Foreign, $15.75).
Current Numbers: Review or Journal, $1.00; **Bulletin,** $.60; **Index,** $2.00.
Psychological Monographs: $6.00 per volume (Foreign, $6.30).
Current Issues: prices vary according to size.

Subscriptions, orders, and business communications may be sent direct to
PSYCHOLOGICAL REVIEW COMPANY,
PRINCETON, N. J.

VOL. 35, No. 3 May, 1928

THE PSYCHOLOGICAL REVIEW

THE ABSOLUTE ZERO IN INTELLIGENCE MEASUREMENT

BY L. L. THURSTONE

The University of Chicago

The object of this paper is to describe a discovery concerning the variability of intelligence by which it is possible to locate its absolute zero. The discovery will be described with four implications, namely, (1) two laws of variability of intelligence, (2) determination of the absolute zero of test intelligence, (3) the construction of a true mental growth curve which has not hitherto been possible, and (4) determination of the age at which test intelligence begins.

The discovery was made by studying a variety of data by means of absolute scaling.[1] The details of this method have been previously described[2] so that it will suffice here to review merely the underlying idea in absolute scale construction. We are here primarily concerned with several of its applications.

The provocation for the absolute scaling method lies in the two most fundamental limitations of intelligence measurement, namely, (1) that we have had no satisfactory unit of measurement for it, and (2) that we have had no origin of measurement except the arbitrary zero score of each

[1] The writer wishes to acknowledge the statistical assistance of Miss Annette McBroom who has carried out most of the calculations for this study.

[2] THURSTONE, L. L., 'A Method of Scaling Psychological and Educational Tests,' *J. Educ. Psychol.*, 1925, 16, pp. 433–451. Cyril Burt's data on the Binet tests were used to illustrate the method. An application to educational scale construction is shown in 'The Unit of Measurement in Educational Scales' by L. L. THURSTONE, *J. Educ. Psychol.*, Nov., 1927. A refinement in statistical procedure for the method is described in 'Scale Construction with Weighted Observations' by L. L. THURSTONE, also to be published in the same journal.

educational and psychological test. The whole study of intelligence measurement can hardly have two more fundamental difficulties than the lack of a unit of measurement and the lack of an origin from which to measure!

If an intelligence test such as the National is given to a large sample of eight-year-old children, we find of course that some of them attain scores much higher than others. In so far as the test is regarded as an adequate index of the function to be measured, we are justified in arranging the children in *rank order* for the trait in question. But measurement implies more. We have no direct means of knowing that the difference in ability represented by the raw scores of 10 and 20, for example, is the same as that between the raw scores of 90 and 100. If the frequency distribution of raw scores is normal *and if we assume that the 'actual' distribution of the trait in question is also normal*, then the difference 20–10 in raw score represents the same increment of ability as the difference 100–90. But, strictly speaking, there is no 'actual' distribution to which we can refer for a final verdict. All that we can ask is that the unit be consistent with a constant meaning throughout the scale. The increments provided by the raw scores do not necessarily satisfy this criterion, because they represent different tasks of different degrees of difficulty.

Another way of stating the insecurity of the raw score unit is that if we were to select a group of children so as to have intentionally a skewed distribution of intelligence, by general impression or otherwise, we could readily select the elements for a test in such a way that the final distribution of scores for this group of children would be normal. Hence the normality of the distribution of raw scores simply proves that the test has been so constructed that the distribution will be normal! It proves nothing regarding the shape of the 'true' distribution, which is in any case indeterminate.

Absolute Scale Construction

Absolute scale construction consists first of a statistical criterion to determine whether two adjacent age groups can

be represented as normal distributions on the same base line. If this criterion is not satisfied, it is impossible to construct a scale so that the overlapping distributions will both be normal on the same scale. Any function of the raw scores that gives a normal distribution for one of these two groups will then necessarily give a skewed distribution for the other. If the criterion is satisfied, then it is possible to find such a function of the raw score that both of the overlapping distributions become normal. This is the manner in which the absolute scale is defined, although the same logic may be applied to try frequency distributions of any form. In absolute scale construction the only necessary assumption is that the distributions of ability in the successive age or grade groups have *the same form*. It is not necessarily assumed that they are normal.

The result of absolute scaling, applicable only when the data satisfy the criterion, is a series of overlapping normal curves on the same base line. These curves represent the distributions of ability in the overlapping age or grade groups. Two constants are determined for each curve, namely, its mean and its standard deviation. The mean ability of one of the groups is arbitrarily chosen as an assumed origin, and its standard deviation on the absolute scale is chosen as a unit of measurement for the whole scale.

The final test of the validity of the construction is as follows. Calculate the scale value of any raw score, or of any test element, with the above unit and origin, and locate it on the scale. Draw a vertical line through this point. Now the proportion of each distribution that lies above this point should agree with the experimentally determined proportion of children in each age group who actually exceed that score or who actually pass that test element. If this final criterion is satisfied throughout the whole range of the scale, then we have a scale construction which is internally consistent with the data throughout the whole range. I have shown elsewhere [3] that this unit of measurement gives results consistent with both psychological and educational test data and that

[3] THURSTONE, L. L., *operibus citatis*.

the generalized 'p.e.' unit which is now in common use gives results that are grossly inconsistent with the data. In the absolute scaling method we have a consistent unit of measurement but the measurement is made from the mean ability of one of the overlapping groups as an assumed origin. Our present problem concerns the discovery of the absolute zero for mental measurement.

Linear Relation between Variability and Mean Test Performance

The discovery is that *a linear relation exists between the standard deviation* (σ) *and the mean test performance* (M) *for the several age groups.* This law holds for a number of psychological tests when their frequency distributions are restated in terms of absolute scaling. It is not universally true for raw scores because of the fact, previously noted, that the raw scoring unit cannot be relied upon to represent a constant increment of ability throughout the whole range.

The generalization made above must be stated with one condition, however, which is implied in the test data from which it is derived, namely, that the social and intellectual factors of selection must operate more or less uniformly for the several age groups. This can be seen if we consider an extreme case. Suppose that we were to select one thousand children of each age for purposes of standardization and that we selected the fifteen year olds exclusively from the eighth grade, while the fourteen year olds were selected from all grades and from the high school. Naturally our norms for the fifteen year olds would not include the brightest children of this age who have already passed the eighth grade nor would we have the dullest children of this age who have not reached the eighth grade. The result would show the variability of the fifteen year olds to be too small in comparison with that of the fourteen year olds even though their respective means were approximately correct. Hence, since the mean and the variability of successive age groups are markedly affected by the conditions of selection of each age group, we must insert the condition of uniform selection on

our generalization. The discovery, so stated, then takes the following form. *With uniform conditions of selection there is a linear relation between the absolute variability and the mean test performance of successive age groups.* This statement refers of course to absolute scaling and not to raw scores.

The above law has been demonstrated for the following tests: National Intelligence Tests; Merrill-Palmer (Stutsman) tests for preschool children; Illinois General Scale; Dearborn Test, Series I; Dearborn Series II; Otis Advanced Intelligence Test; Binet tests (Cyril Burt standardization). The available data for all seven of these tests show a linear relation between absolute variability and mean test performance.

Table 1 is a summary of the absolute variability and the mean test performance of each age group for each of the seven psychological tests. The National test [4] is represented graphically in Fig. 1 where σ is plotted against M. Note the linearity of this graph. Even a physicist could hardly ask for a clearer case of linearity. The Stutsman tests [5] are represented in Fig. 2 in the same manner. Here the errors of measurement are greater because her tests were standardized on only fifty children of each age group, but the linearity is rather clearly indicated. In Fig. 3 the Illinois General Scale [6] has been shown in the same manner, and the linearity is here also pretty clear. The Dearborn tests,[7] Series I and II, are shown in Figs. 4 and 5 in the same manner indicating very clearly a linear relation between variability and mean performance. The Otis Advanced test [8] is represented in Fig. 6 with the same kind of plot. Here it would

[4] 'The National Intelligence Tests,' A manual of directions, published 1924 by the World Book Company. The absolute scaling was done by C. L. Odom for a master's thesis, not yet published.

[5] Data for the Stutsman Test were obtained from Miss Stutsman. They will be published in her doctor's thesis.

[6] Data for the Illinois General Scale were obtained from C. L. Odom's master's thesis, not yet published.

[7] Data for the two Dearborn tests, Series I and II were obtained from C. L. Odom's master's thesis, not yet published.

[8] 'Otis Intelligence Scale,' a manual of directions published by the World Book Company. The absolute scaling was done by C. L. Odom for a master's thesis, not yet published.

be possible to interpret the diagram either as linear or as slightly positively accelerated but Otis himself discards his data for the lowest and for the highest age groups and gives estimated norms for the age groups at both ends of his age range. I have included here only those distributions which he himself uses in establishing age norms for his test. The plot is easily interpreted as linear. In Fig. 7 a similar diagram has been drawn for the Binet test [9] according to the data of Cyril Burt.

The law can be stated in simple algebraic form as follows:

$$\sigma = s \cdot M + K,$$

in which σ is the absolute standard deviation of ability in any age group in terms of the standard deviation of one of the age groups chosen as a base; M is the mean test performance of the same age group, measured in terms of the above unit and from the mean performance of the basic age group as an origin; s is the slope of the graph and K is the intercept on the σ-axis. These two constants will be used for the calculation of absolute zero of test intelligence.

The tests here represented all support the generalization that with uniform conditions of selecting the successive age groups there is a linear relation between the variability of test performance and the mean performance. When the relation of σ and M is plotted for successive *grade* groups, the generalization does not necessarily hold. This is probably because of the progressive selective factors which sometimes govern grade classification.

The Absolute Zero

Perhaps the most interesting application of the linearity just noted is the possibility of locating the absolute zero of test intelligence. Consider first the increase in mean raw score for increasing age. When the mean score is tabulated for each age group we find of course that the mean increases with age at least up to the adolescent years. But the zero score of any one of these tests, the National for example, is

[9] Thurstone, L. L., *operibus citatis.*

not intended to represent zero intelligence. A child may be too young to attain any score above zero on this test but we do not therefore say that its intelligence is actually zero. The zero score in every psychological and educational test is an assumed or arbitrary origin. If we could imagine the scale continuing downward to negative values the very young child might be represented with a certain negative score which would however represent some mental development above the absolute zero intelligence of an inanimate object.

For some time I have considered this problem of an absolute zero as insoluble because no matter how far down we go on the scale of negative scores there seemed to be no *a priori* reason why some organism might not be found with even less intelligence. As long as we deal with raw scores or immediate derivatives of raw scores we really cannot talk about negative values because the scale of raw scores may have no meaning for negative values. As long as the scoring unit is the number of questions correctly answered, or some function of it, the unit is not so defined that any interpretation can be given to negative values. What might be the interpretation of the ability to solve 'minus fifty of these questions'?

Having a unit of measurement the consistency of which can be experimentally checked throughout the whole range of ability we discover the linear relation which has been described. We can now locate the absolute zero indirectly. We have seen that as long as we follow the scale of raw scores downward there is no certain meaning to negative raw scores. If we follow the mean performances, M, on the absolute scale below our assumed origin, there seems to be no direct way of identifying the absolute zero for, here again, how can we know but that some organism might be found whose intelligence would be represented at a point on the absolute scale lower than any point that we might choose?

We arrive at the absolute zero by an indirect route. *The absolute variability of test intelligence must be zero when the mean test performance is absolute zero because, in the nature of the case, the variability cannot be negative.* The procedure of

locating absolute zero consists merely in extrapolating the linear relation of σ and M to ascertain the numerical value of the mean test performance at which the variability vanishes. That value of M will necessarily be the absolute zero.

My original plan was to express variability, σ, as a function of mean test performance, M, and then to find by extrapolation the mean test performance at which the variability vanishes. Much to my surprize I discovered that this relation of σ to M is linear. The only exception that I have so far discovered is the situation in which the variability, when plotted against age, is so erratic that no function can be made out at all.[10]

The linearity of the relation σ and M simplifies our problem of locating the absolute zero because extrapolation of a linear function is easy. We simply continue the linear graph so as to determine the value of M at which the variability vanishes and we have at that value of M the absolute zero that we are seeking. Absolute scale values below that point would require negative variability if the linearity of the function can be assumed to continue throughout the whole range of the scale. Since negative variability would be meaningless we assume that the absolute zero of the scale has been found. We shall present a rather convincing check on this calculation because we may determine by extrapolation of the mental growth curve the age at which test intelligence has a value of absolute zero, and it will be found that this age turns out to be approximately at birth or shortly before.

We may now determine the numerical value of the absolute zero for each of the tests that we have considered. In Fig. 1 we have the linear plot σ against M for the National Intelligence Test. It should be noted that the method of least squares is not applicable to this linear plot because this method, in its usual form, assumes that all of the errors of measurement are in the dependent variable and that the

[10] Such a case is the series of age norms for the Army Alpha test, published by the Bureau of Educational Measurements and Standards, Kansas State Teachers College, Emporia, Kansas, Feb. 1, 1926.

independent variable can be regarded as free from such errors. But here we have errors of measurement in both σ and M. Hence the line was fitted by the two conditions (1) that it should pass through the center of gravity of the plot, and (2) that the slope should be the ratio of the dispersions of σ and of M. I have elsewhere described this procedure of fitting a straight line on psychological data so that its detail will not here be repeated.[11] The points were given equal weight in this study, although in a very refined procedure they can be weighted. By means of the equation

$$\sigma = .2006M + 1.034$$

we determine the value of M for which σ vanishes. It is

$$\text{Absolute zero} = -5.15\sigma_8.$$

In other words, the absolute zero of the ability measured by the National Intelligence Test is about $5\sigma_8$ below the mean

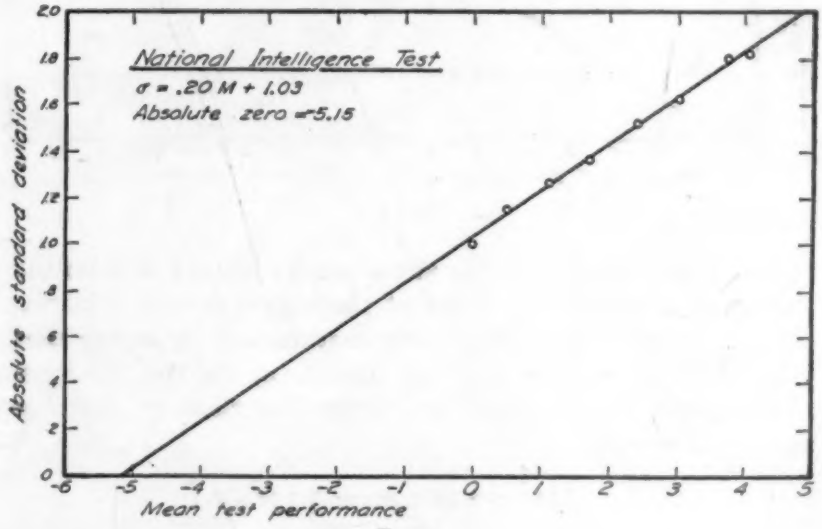

FIG. 1.

performance of the eight-year-old children where σ_8 is the absolute standard deviation of the distribution of this ability in eight-year-old children. If we should represent this fact

[11] THURSTONE, L. L., *operibus citatis*.

graphically we should plot a single normal frequency distribution for the National Intelligence Test to represent eight-year-old children and we should locate the absolute zero at the point -5σ on that diagram.

But here another interesting generalization appears. By Fig. 1 it appears that the absolute standard deviation is directly proportional to the mean performance when the

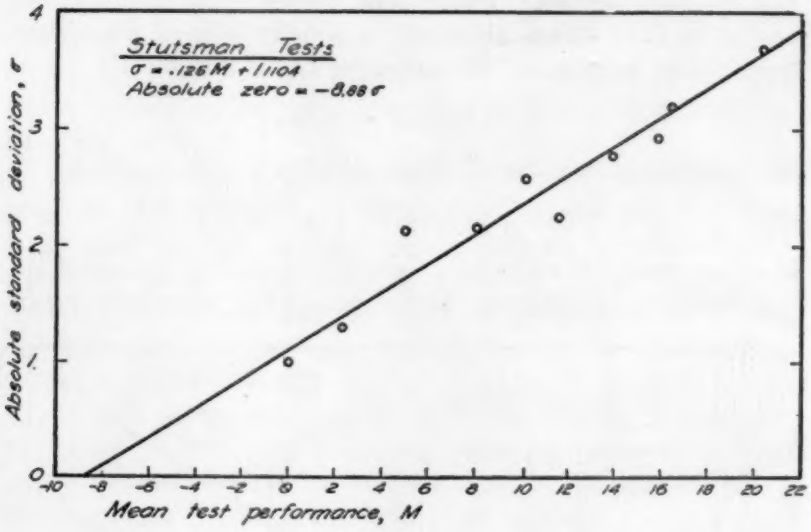

FIG. 2.

latter is measured from the absolute zero instead of from the assumed origin at the mean of the eight-year-old children. Hence we can generalize the above statement by saying that the absolute zero of any age group, n, on the National Intelligence Test is about $5\sigma_n$ below the mean of that age group, or simply,

$$\text{Absolute zero} = -5.15\sigma_n.$$

The higher the mean test performance the greater is the absolute variability of the group and hence the above equation locates the absolute zero as a scale distance below the mean of the group in terms of the dispersion of the age group.

The same analysis is shown for the other tests in the several diagrams. The absolute zero for the Stutsman test

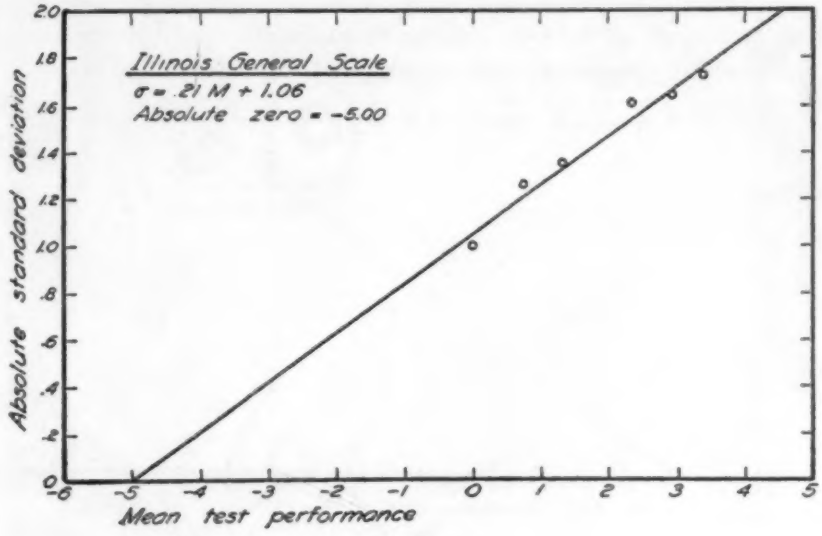

FIG. 3.

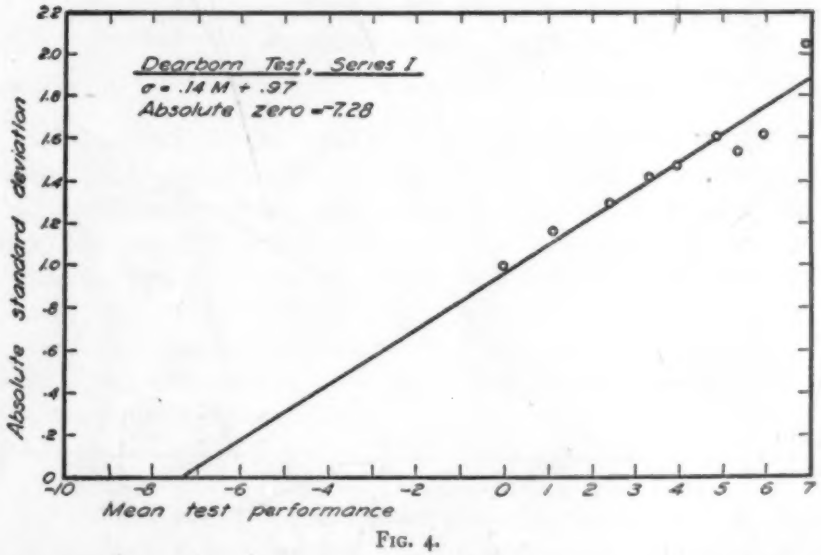

FIG. 4.

is shown in Fig. 2, for the Illinois General Intelligence Test in Fig. 3, for the Dearborn Series I and Series II in Figs. 4

and 5 respectively, for the Otis Advanced Examination in Fig. 6, and for the Binet Test (Cyril Burt) in Fig. 7. In each

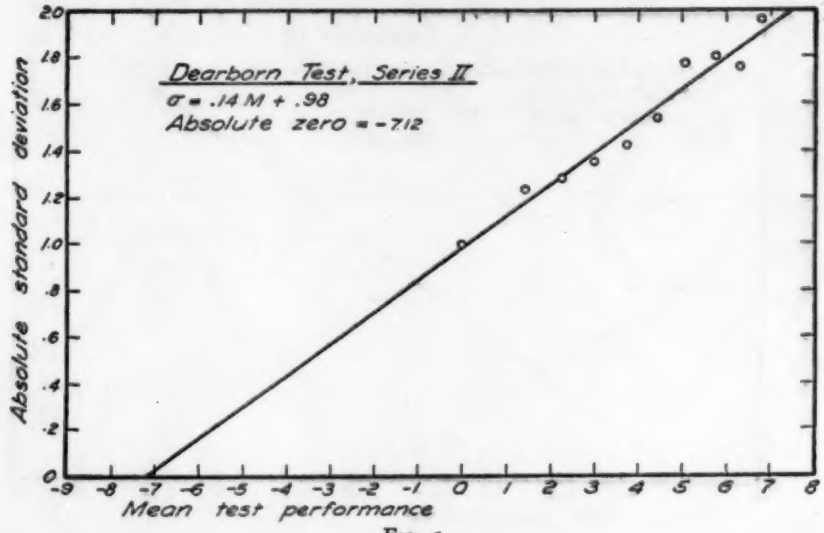

FIG. 5.

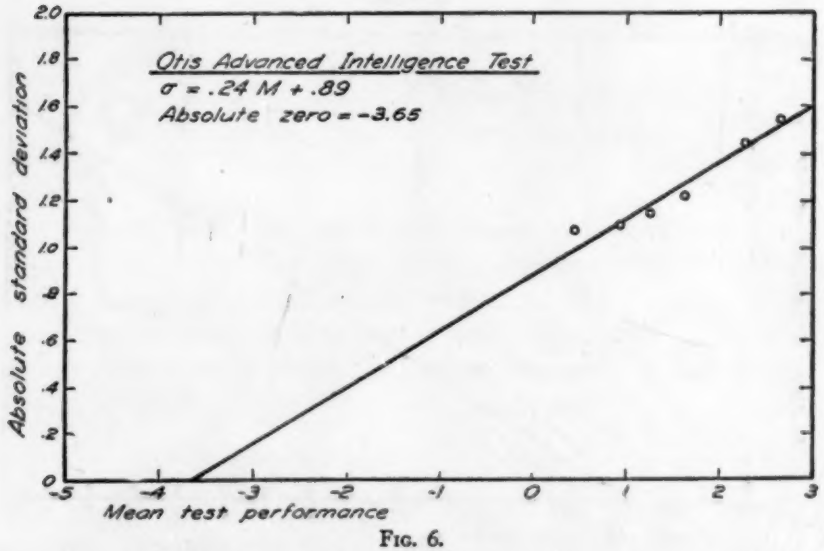

FIG. 6.

of these diagrams inspection first showed the plot to be linear, the equation was then calculated, and the absolute zero determined.

The absolute zero may also be located in terms of the raw score of a test if one is willing to allow an imaginal extension of the raw scores downward into negative values. This may be logically a questionable procedure because there is usually a slightly curvilinear relation between raw scores and absolute

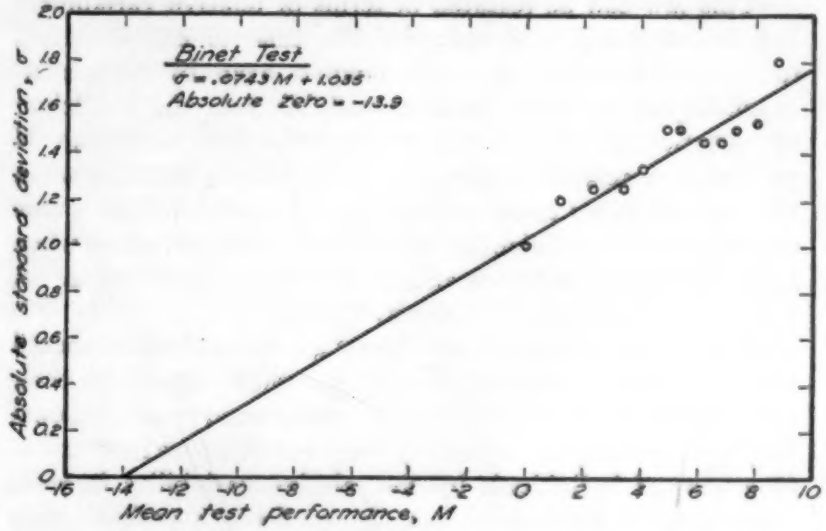

FIG. 7.

scale values. If the relation is plotted and some function established empirically, and if we are willing to assume a fictitious scale of negative raw scores, then it is possible to calculate that negative raw score which corresponds to the absolute zero. Such a procedure is at best rather far-fetched and it will therefore be best to locate the absolute zero as an absolute scale distance below the mean ability of any given age group in terms of the standard deviation of that group as a unit of measurement.

A LAW OF VARIABILITY

Another result of the linearity of the function σ against M is that if we measure M, the mean test performance, from its absolute zero we find that the dispersion is directly proportional to the absolute mean performance. This leads

to the formulation of the following law: *With uniform conditions of selection, the absolute variability in the test intelligence of different age groups is proportional to their absolute mean test intelligence.* This is apparent in all the diagrams where σ is plotted against M for different age groups.

This law can be restated in terms of relative variability, but before doing so it may be advisable to call attention to a very frequent misunderstanding even by prominent psychologists regarding relative variability or the coefficient of variation as it is sometimes called. The coefficient of variation, or relative variability, has a definite meaning when the measurements to which it is applied are made from a true origin but the coefficient of variation is sheer nonsense when it is applied to measurements from an arbitrary origin. A few examples will serve to clarify this difference.

If we measure stature of different groups of men, we should be justified to compare the groups with regard to their relative variability because stature is measured from a rational origin of zero length. Similarly we might compare individuals or groups of individuals by the relative variability in their reaction time because here again we have a rational zero point from which to measure, namely, a zero time interval. But the case is quite different with practically all psychological and educational tests which are scored from an arbitrary zero point. Suppose that two groups A and B have the following constants for their raw scores.

	A	B
Mean Score	50	100
σ	10	20
coef. var.	.20	.20

Here the naïve calculation of the coefficient of variation would show the relative variability to be the same for the two groups. Let the test items be graded in difficulty in this test. We may then assume that both groups passed successfully the five or ten easiest items in the test. The test would differentiate the two groups equally well if these easiest ten items were eliminated or, for that matter, if the author had happened to begin even farther down with ten

items still easier. Now notice that if the author had happened to start the test farther down so that the gross score averages had been, say, 70 and 120, the variability probably would not have been affected because of the addition of twenty very easy questions. But now the coefficients of variation 10/70 and 20/120 would no longer be the same. In other words, the coefficient of variation is affected by the degree of difficulty at which the test itself begins and that is not what we are after when we want to compare the variabilities of two groups. We might just as well compare the relative variabilities of two groups of men as to stature by measuring their heights from a point one foot above the floor! The absolute variabilities of the two groups would remain unaffected but the coefficients of variation would fluctuate depending on where the measurements begin.

In order for the coefficient of variation (relative variability) to have any meaning we must have our measurements from a true or rational origin. The frequent use of the coefficient of variation on psychological test data in terms of raw scores is meaningless. It seems strange that so absurd a procedure can be so common. The absurdity has been well stated by Franzen [12] with special reference to psychological test data.

Since we are here dealing with an absolute zero we are justified in stating the above law of variability in relative terms. The law then takes the following form: *With uniform conditions of selection, the relative variability of absolute test intelligence of different age groups is constant.* It should be evident that this law does not concern the variability of raw scores.

The Mental Growth Curve

In previous articles [13] I have described the possibility of drawing a mental growth curve. At that time I had a rational unit by means of absolute scaling but I did not have an absolute origin. I used as an assumed origin the mean Binet test performance of three year old children. Since we

[12] Franzen, R., 'Statistical Issues,' *J. Educ. Psychol.*, 1924, 15, pp. 367–382.
[13] Thurstone, L. L., *operibus citatis.*

now have not only a rational unit of measurement but also an absolute origin it becomes possible to draw a bona fide mental growth curve. The only representation of mental growth that has so far been possible is the curve of progressive increase in raw score with increase in age. Such curves are of course in common use as norms of performance. It has not been possible to study the function because of two defects, namely, (1) that the unit of raw score gives no assurance of representing equal increments of mental development in different parts of the scale and (2) that there has not been available a true origin from which to measure. Of these two defects the first is by far the more serious, because the nature of a function can be determined even though the origin of measurement be arbitrary, but the nature of the function is itself lost if the unit of measurement is subject to progressive variation of unknown nature.

The mental growth curve which I drew for the Binet test was a legitimate mental growth curve even though the origin was arbitrary, but the function can now be drawn more satisfactorily since we can show it in its true elevation above absolute zero. This has been shown in Fig. 11.

At What Age Does Test Intelligence Begin?

Since we have discovered a method of determining the absolute zero of test intelligence and since we have a consistent unit of measurement for test performance, it becomes possible to draw a mental growth curve, and the question naturally arises as to whether this mental growth curve passes through absolute zero at or near birth. It should be noted that in determining the absolute zero we have not made use of chronological age measurement. If it should turn out that the absolute mental growth curve passes through the absolute zero at some preposterous age such as minus five years or plus several years, then we should be suspicious of the method by which the absolute zero was determined. It is just this sort of common sense check to which we can submit the location of absolute zero. We should expect the mental growth curve to pass through the absolute zero at or near birth, or not more than nine months before.

For the purpose of this practical check we shall use four tests, namely the National, the Stutsman, the Otis Advanced and the Binet as standardized by Cyril Burt in London. We shall use the Stutsman test for preschool children although standardized on small groups because her norms extend down to the age of 21 months and consequently the extrapolation for that test covers a relatively short age range. This is a decided advantage because if the norms begin at eight years it is necessary to extrapolate the mental growth curve over an age range practically equal to the range of the whole standardization from 8 to 16 and this introduces of necessity a possible error of several months in the age at which the growth curve passes through absolute zero. This error is relatively smaller when the extrapolation covers only a short age range such as eighteen months.

We shall start with the National Intelligence Test. In Fig. 8 the mean test performance is plotted against chronological age with data obtained directly from Table 1. The growth curve seems to be linear over the age range studied. This may be due to intrinsic linearity of the mental growth curve, but it is more likely that a progressive selective factor is superimposed on the natural negative acceleration of the mental growth curve so as to give it the appearance of linearity. However, since the data show the curve to be linear we shall treat it that way. We draw a horizontal line at the level of absolute zero which has been determined in Fig. 1. The linear growth curve is extended downward until it intersects the level of absolute zero. We then find that the age at which the mental growth curve passes through absolute zero is zero age, namely at birth! This determination is accurate at least within several months if the continued linearity of the growth curve can be assumed.

At this point a word should be said regarding the probable shape of the mental growth curve near its origin. It is hardly likely that the growth curve starts with a negative acceleration. It is more likely, and in fact more natural, to expect it to begin with a positive acceleration and to change gradually to a negative acceleration which continues toward maturity.

TABLE I

Mean age in yrs.	Binet (Burts' Data)		National Intelligence Test		Dearborn Series I		Dearborn Series II		Illinois General		Otis Advanced		Mean age in mos.	Stutsman Test	
	M	σ	*M*	σ	*M*	σ	*M*	σ	*M*	σ	*M*	σ		*M*	σ
3.5	0.000	1.000											21 mos.	0.0000	1.0000
4.5	1.231	1.191											27	2.2222	1.3072
5.5	2.336	1.255			0.000	1.000							33	4.9858	2.1363
6.5	3.372	1.255			1.072	1.165							39	8.0386	2.1642
7.5	4.061	1.333			2.394	1.292	0.000	1.000					45	10.1406	2.5770
8.5	4.875	1.496	0.000	1.000	3.315	1.415	1.369	1.231	0.000	1.000			51	11.5982	2.2283
9.5	5.433	1.500	0.512	1.162	3.955	1.485	2.241	1.276	0.737	1.258			57	13.9082	2.7822
10.5	6.190	1.453	1.082	1.269	4.849	1.606	2.990	1.357	1.315	1.362	0.440	1.068	63	15.7360	2.9286
11.5	6.776	1.446	1.711	1.371	5.321	1.554	3.748	1.425	2.368	1.621	0.941	1.092	69	16.3761	3.2007
12.5	7.314	1.503	2.418	1.533	5.944	1.622	4.402	1.553	2.929	1.645	1.236	1.154	75	20.3180	3.6947
13.5	8.040	1.532	3.028	1.633	6.890	2.050	5.062	1.774	3.382	1.728	1.606	1.222			
14.5	8.684	1.792	3.761	1.806			5.734	1.798			2.240	1.441			
15.5			4.040	1.818			6.329	1.757			2.627	1.552			
16.5							6.767	1.959							

It is probable that the mental growth curve should be represented as theoretically asymptotic to the level of absolute zero so that the sperm and the ovum would represent not zero but infinitesimal amounts of the process which matures into measurable intelligence. The mental growth curve may ultimately be shown to have the characteristics of Raymond Pearl's generalized population growth curve.[14]

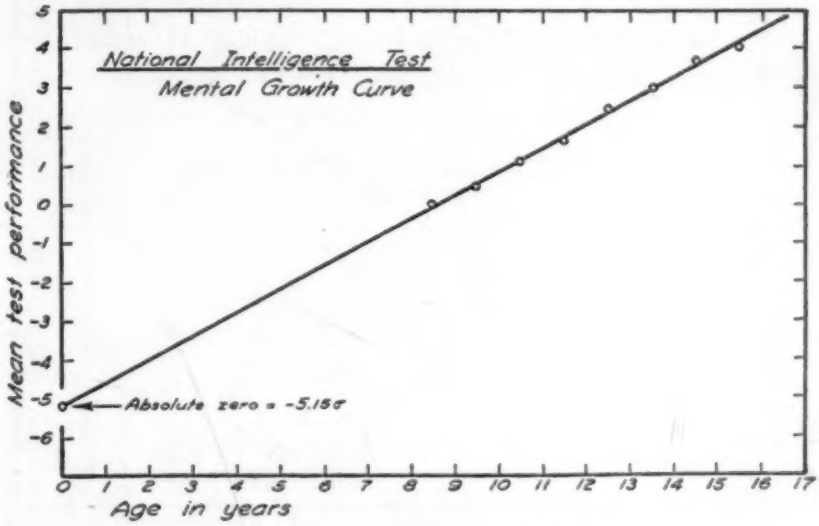

FIG. 8.

The continuity of the growth curve in Fig. 8 through the absolute zero at birth and the mean performances at the several age levels is very striking. The linear growth curve of Fig. 8 as determined by the eight means from the data can be projected downwards until it reaches zero age and the value of the mean at that point agrees with the absolute zero as determined by the variability diagram of Fig. 1. This is a common sense test of the reasonableness of our determination of absolute zero.

The growth curve for the Stutsman test is shown in Fig. 9.

[14] This reasoning is in agreement with Professor Culler's analysis of the learning function. For another discussion of initial positive acceleration of the learning function, see my monograph 'The Learning Curve Equation,' PSYCHOL. MONOG., 1919, 26 (No. 114), pp. 51.

The ten mean performances were first plotted from Table 1. The absolute zero was determined in Fig. 2. Now when the growth curve of Fig. 9 is projected downward it is found that it reaches the level of absolute zero at birth. This is again rather striking evidence of the reasonableness of the determination of absolute zero. The graph in Fig. 9 could be

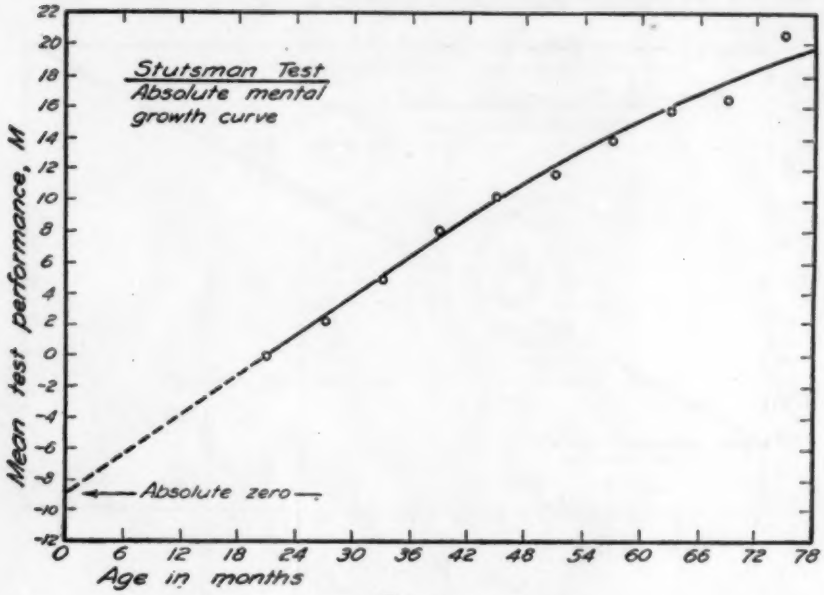

FIG. 9.

represented as a straight line. That has been done and the equation calculated as previously described. Such a line reaches the level of absolute zero at an age of about minus three months, *i.e.* about three months before birth. It is probable that absolute scaling of the developmental reflexes in young infants would reveal a growth curve of positive acceleration which begins with the early development of the foetus. The choice between a straight line and a line of slight curvature as shown in Fig. 9 is not here of primary importance because in either case the growth curve is at the level of absolute zero at or slightly before birth.

The mental growth curve for the Otis Advanced examination is shown in Fig. 10. It was plotted from the data of

Table 1. The relation seems to be linear and when it is extended it reaches the level of absolute zero at birth. The fact that these mental growth curves when extrapolated pass through absolute zero at birth is rather convincing evidence that the mental functions which are operative in the adult intelligence tests begin their development at birth even though these functions cannot be measured at a very early age. It might possibly seem as though the functions that are operative in taking the paper-pencil tests in group form begin to develop at the time when the writing coördination

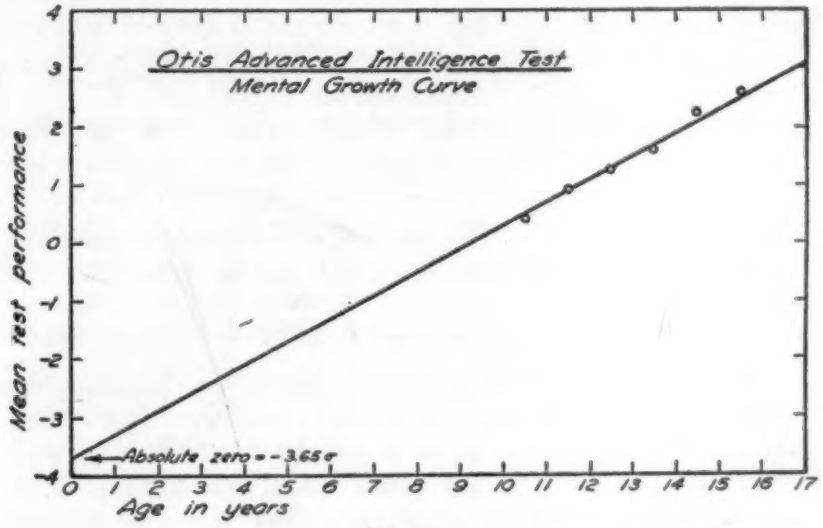

FIG. 10.

is begun, but these results show rather definitely that whatever it is that is measured by the group intelligence tests, it begins its development at birth or shortly before. This is a psychological finding of no mean interest.

The mental growth curve for the Binet test is shown in Fig. 11. The twelve points on this curve are again plotted from Table 1. It appears immediately that this growth curve is not linear. It has a negative acceleration and it can readily be thought of as continuous with the absolute zero of performance at birth. The continuity of this function is rather striking. Owing to the curvature of this function the

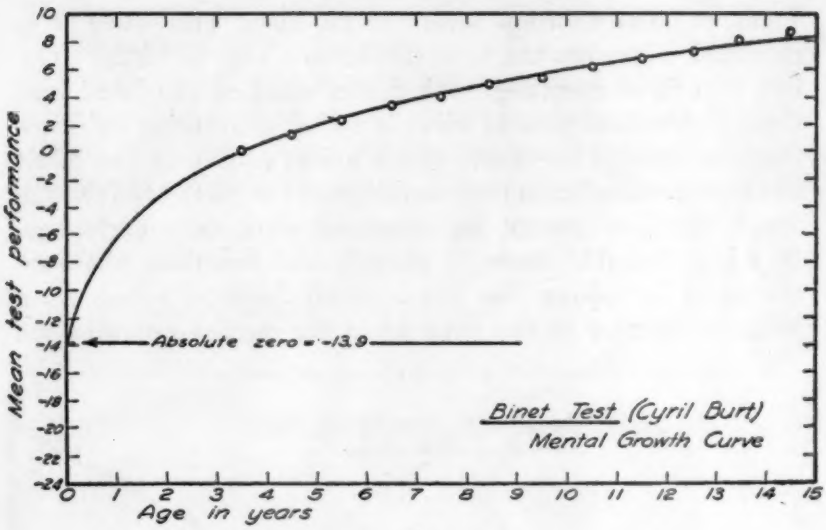

FIG. 11.

determination of age at absolute zero is not nearly so accurate as in the other three tests for which mental growth curves are shown.

SUMMARY

1. The discovery which has been verified so far on seven psychological tests is that with uniform conditions of selection there is a linear relation between the absolute variability and the mean test performance of successive age groups. This generalization refers to absolute scaling and not to raw scores.

2. The absolute zero is located indirectly. The absolute variability of test intelligence must be zero when the mean test performance is absolute zero because, in the nature of the case, the variability cannot be negative. The absolute zero is located by extrapolating the above linear relation to ascertain the scale value of the mean performance at which the variability vanishes. That scale value is the absolute zero. It is defined as a distance below the mean performance of any age group in terms of its own standard deviation.

3. Having found the linear relation above described and having located the absolute zero, the following law of vari-

ability is a necessary inference: With uniform conditions of selection, the absolute variability in the test intelligence of different age groups is proportional to their absolute mean test intelligence. This law can be stated in terms of relative variability as follows: With uniform conditions of selection, the relative variability of absolute test intelligence of different age groups is constant. These laws refer to absolute scaling and not in any sense to raw scores.

4. By means of the unit of measurement provided by absolute scaling and the absolute zero it becomes possible to construct a true mental growth curve for a specified mental test. It has not hitherto been possible to study the function of mental growth because of the lack of a unit and an origin.

5. The validity of the determination of absolute zero is subjected to a practical test by determining the age at which the mental growth curve passes through absolute zero. It is found that this happens at birth or shortly before.

6. The fact that the mental growth curve passes through absolute zero at or before birth constitutes statistical evidence that test intelligence begins its development at this early age even though it is not then directly accessible for measurement.

[MS. received November 25, 1927].

INTELLIGENCE AS THE CAPACITY FOR VARIABILITY OR VERSATILITY OF RESPONSE[1]

BY A. S. EDWARDS

University of Georgia

It is a question whether the term intelligence can be used in systematic psychology in a scientific sense. Its use has been essentially practical and the term has been taken over from popular use. As a technological term its place is undeniable and various interpretations of intelligence for practical purposes are not to be questioned if they are of practical value. Our present interest is to consider a definition that may possibly put intelligence in its place in systematic psychology, if it has such a place. It may eventually be better to use another term that shall be a definitely scientific one. We can, however, study the concept of intelligence without going out of our way at this time to propose the new term.

We believe that the intelligence increases as we rise in the scale from the lowest plant life to the highest animal life and that it finds its largest measure in the superior adult.[2] What characteristics do we find increasing with the increasing intelligence? Let us consider some of these organisms.

The fern outside my window carries on the vital processes; it can stretch up and out for a little distance but it is confined to a very small number of responses, practically those of individual living and of reproduction. The moonvine on the back porch adds to these performances the ability to climb up the side of the porch and house; the sensitive plant closes when it is touched and a number of flowers close upon their victims to obtain food; roots of trees may

[1] Address of the president, Southern Society for Philosophy and Psychology. University of Georgia, April 22, 1927.

[2] Jennings, H. S., 'Behavior of the Lower Organisms,' Macmillan, 1906, espec. 334*f*; Terman, L. M., 'The Measurement of Intelligence,' Houghton Mifflin Co., 1916; Washburn, M., 'The Animal Mind,' Macmillan.

go deep or enter water pipes; but when we sum up all the varieties of plant activities we find that responses are of a very limited range. The plant can live only upon what comes within a very narrow range of environment; it cannot move from its station to obtain food or to escape its enemies. Whether or not it has intelligence has been a moot question.

The lower animal organisms behave differently: they move about in search of food and to escape their enemies. As we rise in the scale of animal life we find animals acting together, playing and building homes together and entering into many activities such as recognizing friends and enemies, gathering food in the autumn and hiding it in safe places, acquiring new tendencies to response when instructed or trained by human beings, etc. But still the activities are within a narrower range than are those of man.

Man presents the greatest range of activities of any living things. He not only builds houses but makes them of many different kinds; he develops languages, sciences, arts, literature, customs, religions; he travels on land and sea, under the surface of the earth, beneath the waves, and girdles the mundane sphere with air-ships and radio. He records the happenings of the past and plans for the future. He modifies himself to fit various environments, and—of the greatest importance—he changes environments to suit himself.

As we thus consider the differences among animals of different degrees of intelligence we can in the same way consider human beings of different degrees of intelligence. In the class we call idiots we find ability to do very little that we call distinctly human. The idiot cannot even feed himself and attend to his common wants. With more intelligence an individual feeds himself though he may eat anything—an idiot may need a new shirt nearly every morning because he has eaten too much of the one worn the day before. One having what we call higher intelligence is said to 'eat discriminatingly,' and can do a little to take care of himself. As we rise farther and farther in the scale of intelligence we find that the individual can play, that he

can work a little, do more and more complicated things, work without direction and oversight, and then that he can plan and execute original and very complicated designs.

How can we define this something we call intelligence which we find to increase as we rise in the scale of animal life and as we pass from the idiot to the supernormal individual? The intelligence of psychologists has not left us wanting in variety of definitions.[3] It is a means of adaptation; a means of adjusting to new situations; a means of success; ability to learn; ability to think to good effect, to appreciate situations and easily find solutions, to appreciate ends and to be able to work consistently toward those ends, the power of good responses from the point of view of truth or fact; ability to carry on abstract thinking; intellect plus knowledge; general capacity for independence, originality, and productiveness in thinking; a mechanism for integrating and giving a somewhat unified effect in behavior; capacity to inhibit instinctive adjustment, to readjust in terms of imaginally experienced trial and error, plus volitional capacity to act to greater advantage to the individual; capacity to acquire capacity; a group of mental processes: sensation, perception, association, memory, imagination, discrimination, judgment and reasoning, etc.

If variety is the spice of psychology the student cannot complain that he fails to find a rich diet when he attempts to digest the subject of intelligence, to say nothing of the feasts prepared for him in the 'Psychologies of 1925' and those of other years. A little indigestion is only to be expected; if some degree of pessimism result it is at least not worse than too much optimism. Since the custom of bringing contributions to the feast is well established it is not inappropriate for the present writer to bring another definition of intelligence. First, however, we shall give consideration to some of the definitions above mentioned.

Most of the definitions of intelligence have been made from the points of view of philosophy or of applied psychology and the clinic; from the point of view of general psychology,

[3] See especially *J. Educ. Psychol.*, 1921, 12, 123*ff*; 195*ff*; 271*ff*.

we must recognize the value of definitions made from particular points of view even though they may not be acceptable for pure science. Adverse criticisms that follow do not apply equally from the points of view of philosophy and of technology, for definitions may have been of great value for the purposes for which they were made and as hypotheses upon which to work. As such we grant them their full measure of value and consider them now only from the point of view of general psychology and of fact.

For pure science the term adaptation involving teleology is questionable: actions must be evaluated; ends must be recognized; we are in the realm of philosophy or technology. The fact of a correlation between intelligence and adaptation is still more to be questioned. The idiot has more intelligence than plants or protozoa but is not better adapted; a nonconformist or reformer may have the highest intelligence and be less well adapted to his environment than the conformist of mediocre intelligence. It may even be questioned whether human beings are as well adapted to their environments as are the lower animals. Adaptation does not appear to be a satisfactory criterion of intelligence. High intelligence may maladapt as well as adapt to environment.

Success is less satisfactory as a criterion, for we do not know what we mean by success unless we make an arbitrary definition in terms, for example, of some standardized test. But we do not know that the standard test is a test of intelligence and we are dangerously near to travelling in a circle. Even if we assume to know what we are talking about when we use the term success, we are forced to admit that many men of mediocre ability who stay carefully within the field of their training may be adjudged successful, whereas men of the highest intelligence who try many new things, who attempt research and invention, may be considered as failures. Some men fail in one kind of work and succeed in others. A and B grade men, as indicated by intelligence tests, often fail at C and D grade jobs.

As ability to learn, the definition contains much truth but it is too narrow.[4] The intelligent may learn but they may

[4] Peterson, J., PSYCHOL. REV., 1922, 29, 387*f*.

also unlearn so far as that is possible. What is more important, they may use what they have learned in situations that are not to be called those of learning, such as the making of decisions and of deductions from what has been learned in the past. It might be more accurate to say: Intelligence is the ability to use what has been learned.

Valuable suggestion is found in the consideration of integrative activity and of the unified effect in behavior. But this definition overlooks the fact that the intelligent action may be disjunctive as well as associative;[5] and it does not differentiate the action of the less intelligent or non-intelligent centers such as the medulla and the cerebellum from that of the cerebrum. Animals quite low in the scale of intelligence may act with a greater integration of behavior and a better unified effect than do men of the highest intelligence. Indeed, intelligence with its many inhibitions, its fine discriminations, its hesitations, may interfere directly with unified action. The highly intelligent man may stutter or act with much interference to a unified effect when a man of less intelligence proceeds fluently and easily.

Organization can hardly be posited as a differentia of intelligence, since it exists prior to intelligence and is accepted as a characteristic of non-intelligent activity. It is definitely present in plants and in reflexes and in the non-intelligent activities of the medulla and the autonomic nervous system, and the idea of organization is quite fundamental to the activities of the whole material universe. Our definition does not omit organization but assumes it. And the higher the intelligence the greater the organization, not only passive but active. But organization is not the differentiating factor. In both the intelligent and the non-intelligent activities we find organization; in the non-intelligent we do not find versatility, resourcefulness, the possibility of a large range of response; in the intelligent realm we do. Thus we do not question the importance of organization, emphasized as it has been by Ebbinghaus, Wundt and others, but we question it as being as truly a differentia of intelligence as is versatility.

[5] Sherrington, C. S., 'The Integrative Action of the Nervous System,' 1906, pp. 279–299; 388 *ff.*

Definition in terms of inhibition, readjustment in terms of imaginally experienced trial and error, and volitional capacity to act to greater advantage to the individual, may be useful from a practical point of view; intelligence probably means greater capacity for inhibition and for thinking; but the term *advantage* again involves evaluation and some technological standard; the intrusion of the idea of volition takes us into a realm somewhat foreign to that of intelligence. The latter may often be negatively correlated with volition; the more exact and discriminating a man becomes, the less may he be inclined to act.

Definitions in terms of knowledge, association, and memory seem rather to point away from intelligence than toward it, or to point to the habitual and to the results of earlier intelligent activity than to present intelligence. Memory shows no high positive correlation with intelligence: feeble-minded individuals may have excellent memories, and those of the highest intelligence may have mediocre memories.

Bergson has pointed to the differentia of intelligence as being "the faculty of manufacturing artificial objects, especially tools to make tools, and of indefinitely varying the manufacture."[6] Disregarding his use of the word translated *faculty* (I have not a copy of the original available), Bergson misses the point, for he says on the preceding page, "As regards human intelligence, it has not been sufficiently noted that mechanical invention has been from the first its essential feature." Why he should have hit upon mechanical invention rather than upon the inventions in literature, art, philosophy and scientific theory is hard to understand. The latter are certainly not less intelligent acts. If he had emphasized the idea of indefinite variation of behavior he would have anticipated the main point of this paper.

Definitions in terms of ability to do abstract thinking, general capacity for independence, originality, and productiveness in thinking seem to be more nearly in line with the facts and to avoid the difficulties indicated in the criticisms of other views. It seems possible to come a little closer to the

[6] Bergson, H., 'Creative Evolution,' tr. by A. Mitchell, Henry Holt and Co., 1911. 139

nature of intelligence and to define in such a way as to be true to at least a very large number of the facts in the case. It is proposed to define intelligence as *capacity for variability or versatility of response.*[7]

This definition means that with greater intelligence there is greater capacity to succeed or to fail, to make greater and more successes or failures, to adapt or to maladapt, to learn more things, to make more applications of what has been learned, to make finer distinctions, to spin more theories, to make a larger number of trial and error responses, to act in terms of more ends, including more remote ends, etc. But it must not be forgotten that it means more mistakes as well as more successful responses. The finer intelligence can make more abstractions and generalizations, can be more general and comprehensive, more productive and original, can do more unhabitual things, solve more problems, and find more solutions for the same problems than can the poorer intelligence; the range or scope of analysis and selection is vastly increased. Reflex means relative invariability; intelligence means variability, flexibility, versatility, plasticity, resourcefulness.

For this view we shall now consider certain lines of evidence. 1. The feeble-minded individual who has been trained to certain work may be practically unable to make mistakes; he has not enough intelligence. But the man of more intelligence may make many mistakes. Men of the highest intelligence, indeed, have made statements directly pertinent to the present point.

[7] Versatility or variability as *a* characteristic of intelligence is, of course, not new. As Thorndike told us a long while ago: "The fact is that original tendencies range with respect to the nature of the responses from such as are single, simple, definite, uniform within the individual and only slightly variable amongst individuals, to responses that are highly compound, complex, vague, and variable within one individual's life and amongst individuals . . . . They range with respect to the bond or connection from slight modifiability to great modifiability, and from very close likeness amongst individuals to fairly wide variability." 'Educ. Psychol., Briefer Course,' 1914, 4–5. Our suggestion is not this but that the capacity for versatility is *the* fundamental thing in intelligence and that it appears in richer responses, in finer discriminations, in more ideas of end, in ability to learn more things, and the like.

The term 'versatility' has also been used as, for example, in 'Behaviorism and Psychology,' by A. A. Roback, p. 167.

Jevons in 'The Principles of Science,' pages 577–78, writes as follows:

> It would be an error to suppose that the great discoverer seizes at once upon the truth, or has any unerring method of divining it. In all probability the errors of the great mind exceed in number those of the less vigorous one. Fertility of imagination and abundance of guesses at truth are among the first requisites of discovery; but the erroneous guesses must be many times as numerous as those which prove well founded. The weakest analogies, the most whimsical motions, the most apparently absurd theories, may pass through the teeming brain, and no record remain of more than a hundredth part. There is nothing really absurd except that which proves contrary to logic and experience. The truest theories involve suppositions which are inconceivable, and no limit can really be placed to the freedom of hypothesis.
>
> Kepler is an extraordinary instance to this effect. No minor laws of nature are more firmly established than those which he detected concerning the orbits and motions of planetary masses, and on these empirical laws the theory of gravitation was founded. Did we not learn from his own writings the multitude of errors into which he fell, we might have imagined that he had some special faculty of seizing on the truth. But, as is well known, he was full of chimerical notions; his favorite and long-studied theory was founded on a fanciful analogy between the planetary orbits and the regular solids. His celebrated laws were the outcome of a lifetime of speculation, for the most part vain and groundless. We know this because he had a curious pleasure in dwelling upon erroneous and futile trains of reasoning, which most persons consign to oblivion. But Kepler's name was destined to be immortal, on account of the patience with which he submitted his hypotheses to comparison with observation, the candour with which he acknowledged failure after failure, and the perseverance and ingenuity with which he renewed his attack upon the riddles of nature.
>
> Next after Kepler perhaps Faraday is the physical philosopher who has given us the best insight into the progress of discovery, by recording erroneous as well as successful speculations. The recorded notions, indeed, are probably but a tithe of the fancies which arose in his active brain. As Faraday himself said—"The world little knows how many of the thoughts and theories which have passed through the mind of a scientific investigator, have been

crushed in silence and secrecy by his own severe criticism and adverse examination; that in the most successful instances not a tenth of the suggestions, the hopes, the wishes, the preliminary conclusions have been realized."

. . . It may perhaps be said that Faraday committed to the printing press crude ideas which a friend would have counselled him to keep back. There was occasionally even a wildness and vagueness in his notions, which in a less careful experimentalist would have been fatal to the attainment of truth (579).

From the foregoing it appears that intelligence is not something that keeps one from errors, mistakes, inaccuracies, but that actually intelligence makes more mistakes possible and it seems that the greater the intelligence the greater is the capacity for making them as well as for making correct responses. The characteristic that is unmistakable is the ability to respond in more ways, to make more theories, to invent more ways of solving problems, *i.e.*, capacity for variability of response. Outside of applied psychology with its norms for tests, and philosophical evaluation, etc., it is not the province of psychology, that is to say, it is not the business of pure science, to say whether or not the responses are good or bad any more than it is to assume to define reasoning in terms of the correctness or validity of the reasoning.

2. Our definition is comprehensive of the factors involved in intelligence without including those which do not belong to it. It includes ability to learn and also ability to unlearn and to use what is learned; ability to adjust and also to maladjust one's self to the environment and also to modify the environment so as adjust it to the individual; ability to do various kinds of thinking and also ability to react in new ways with or without reasoning. It distinguishes intelligence from habit, memory, and volition which do not universally have high positive correlation with intelligence.

3. It avoids the assumptions of teleology which takes us out of the realm of science into philosophy; we are not thus called upon to evaluate acts as good or bad, adaptive, successful, etc.

4. It saves us from the doubtful statement that the more intelligent individual is better adapted than the less intelligent and from the false idea that correlates intelligence with success whatever that is.

5. The definition is independent of practical norms and results of the intelligence tests now in use and may give emphasis to the effort to construct tests that shall attempt to discover so far as possible the resourcefulness of individuals in problematic situations.

6. So far as I know, the definition is harmonious with the known facts of biology, of neurology especially, and not in conflict with any of them. The increasing complexity of the nervous system as we rise in the scale of intelligence gives just this greater possibility of response in our environments. The more complicated and highly developed nervous systems in the higher animals and especially in man correlate directly with the larger possible number of connections, associations, inhibitions, in a word, responses.

7. In different systems of psychology the definition in terms of variability or versatility is applicable. In introspective psychology greater intelligence might be found to be attended by a richer ideational content, more ideas of end, recognition of more meanings, etc. Or, the greater intelligence has the possibility of more configurations, or of more relationships with the environment, or of more conflicts, defense reactions, and solutions of conflicts, etc.

8. History will probably corroborate our suggestion. It has been said to be the record of men's blunders. In an historical survey, intelligence will probably be found to be more nearly correlated with variability or versatility of response than with correctness of action, wisdom and foresight. The history of science likely shows more efforts of intelligence cast aside than kept and a larger number of moves into blind alleys than toward the opening out of the maze.

Objections

It may be objected that the insane show more responses than the normal individual and therefore according to our

definition indicate more intelligence. But more kinds of activity on the part of the insane do not indicate more *capacity* for variability or versatility of response; there may simply be less inhibition than in normal cases. Or, diseased or toxic conditions may effect actions that would not occur in their absence and here we have conditions quite independent of the intelligence of the individual.

If it be asked whether variability of emotional response indicates intelligence, the answer is that variability of such emotional response probably depends upon variability of intelligent response.

In view of the controversies of the past it must be expected that the objection will be raised that intelligence cannot be properly defined from the objective point of view and must be defined subjectively in terms of conscious learning, ideas of end, purpose, choice, and the like. If we cling to this type of definition we are forced either to the uncertain method of logical analogy when we consider the intelligence of lower animals or to having two definitions of intelligence, the one for human beings, the other for lower animals. Our definition avoids these difficulties and provides for the subjective factors: versatility of response may be in the richness of ideational activity, in the larger range of ideas of end.

The present definition does not commit us to either theory of intelligence as regards its being general or specific. Objection is hardly possible at this point. In view of the facts we have at present we think of specific kinds of intelligence: literary, mathematical, abstract, concrete.[8]

[8] It is tempting to expand on several of the above topics to an extent which space does not permit. The view of animal intelligence adequate for lower organisms needs to be supplemented for the behavior noted in the human being. If in lower animals intelligent behavior is essentially the making of new associations (Holmes, S. J., 'The Evolution of Animal Intelligence,' 1911, esp. 164), or the modification of behavior in accordance with experience (Jennings, H. S., 'Behavior of the Lower Organisms,' 1906, esp. 334), this conception is not adequate to human intelligence as we have tried to show. It would be interesting to discuss the paper of Dodge on 'Problems of Human Variability' (*Science*, 1924, **59**, 163*ff.*) in this connection but a reference must suffice. As for practical applications, the suggestion for the highest type of teaching is clear: to foster versatility, originality, to encourage idiosyncrasy, to arouse the individual to develop the best of the special talents he has instead of being leveled to the common standard—all of which is old enough in theory but sadly lacking in practice.

It is conceivable that someone might say that the variable action of an engine that jumps the track is by our definition an indication of intelligence; or, the unexpected action of a chemical might be thought of as coming under this conception of intelligent action. Our first answer is that our definition is made in the field of psychology which probably does not include the phenomena mentioned. If these and like unexpected actions are to be included under psychological definition, the reply is that no capacity for variability of response is indicated that is worthy of the name; such variability does not at all compare with the possibilities in the least intelligent of the feeble-minded, and thus if intelligence is indicated it is of the lowest type and approximates zero.

If objection is raised in connection with the possibilities of measurement the answer is simply that the situation is none the more difficult because of the definition. In the realm of the more fixed and less modifiable reflexes and habits, for example, we have the greatest expectation of measurement and prediction; in the realm of the intelligent, the variable and versatile, we have the least expectation of exact measurement.

The point of special emphasis in this paper is not that variability or versatility has not been recognized in connection with intelligence, but rather that it is much more fundamental for our conception of intelligence than has generally been recognized. The idea of simple adaptation breaks down when we reach human responses. Ability to learn likewise is too limited, for it omits the ability to use what has been learned. Modifiability suggests too narrowly the changes within the individual, for the most intelligent acts may result in more modification of others or of the environment than of the self. And holding to ends or purposes may be intelligent or obstinate; the less intelligent may hold rigidly to certain ends for want of ideas; if helped by others of more intelligence to think out ends which had been beyond them, they may shift to new ends and purposes and repeat their obstinacy for these.

The Place of Intelligence in Systematic Psychology

Whether or not the term intelligence shall find a rightful place in systematic psychology is not perhaps to be answered at this time. If it does our treatment suggests a line of development from reflex action to intelligent behavior and is applicable from the intelligent behavior of the lowest organisms in which it is present to that of the most intelligent human being. The least scope or range of reaction is found in the reflex; the largest amount in behavior of the highest species, varying amounts of versatility between. Alike genetically, the dispositions to intelligent behavior become so different in mature development of the human adult that intelligence has even been opposed to innate tendency from certain points of view. From the reflex to the intelligent we have the two extremes with all varying degrees between; in other words, from the most fixed we pass to the most variable or versatile.

Remembering that a reflex can be very considerably modified we may say: a reflex action is one out of one or of a very limited number of possible responses to a stimulus; an intelligent action is one out of an indefinite number of possible responses.

[MS. received December 8, 1927]

THE THEORY OF FACTORS,[1] I

BY STUART C. DODD

National Research Council Fellow 1926–27

In this paper will be reviewed Spearman's theory of two factors, Thomson's sampling theory, and the more general theory of specific, group, and general factors.

In 1904, Spearman published an investigation (1) of the relation between sensory discrimination and intellectual abilities. In his paper, he was one of the first to use correlation coefficients, probable errors, and corrections for attenuation. His measures of the so-called *sensory* type were tests of pitch, brightness, and weight discrimination. The measures of *intellectual ability* were obtained from ratings by teachers and schoolmates. There were three groups of school children and one group of adults, with about thirty subjects in each group. The outstanding result was that the correlation between the pool of his tests of 'general sensitivity' and the pool of his measures of 'general intelligence' was approximately one, when corrected for attenuation. From this, he drew the conclusions: (*a*) "Whenever branches of intellectual activity are at all dissimilar, then their correlations with one another appear wholly due to their being all variously saturated with some common fundamental Function (or group of Functions)" (p. 273); (*b*) "all branches of intellectual activity have in common one fundamental function (or group of functions), whereas the remaining or specific elements of the activity seem in every case to be wholly different from that in all the others" (p. 284); and (*c*) "the specific element can to a great extent be readily eliminated by varying and combining the kind of test"

[1] The author wishes to thank Professor C. Spearman and Professor G. H. Thomson for their willing coöperation in this study and Dr. Garnett and Professor K. Pearson for assistance in difficult points. Thanks are also due to Mr. C. R. Brolyer for verification of the formulas and valuable suggestions regarding the text.

whereby means are afforded "for investigating and diagnosing the central Function" (p. 274). He wished to reserve his description of this fundamental 'Function' until he knew more about its objective relations.

The investigation was weak in the following particulars: (1) the small number of subjects, (2) *estimated* reliability coefficients, (3) the assumption of chance distribution of errors of observation, and (4) the lack of objective measures of intellectual ability. This study was the first quantitative experimental evidence for the existence of general ability underlying both intellectual and—to a lesser extent—sensory proficiencies. In spite of the weaknesses enumerated, this first study was verified by subsequent investigations (**2, 3**).

Spearman's next paper (**2**), written with Krueger, introduced the concept of group factors, as they found at least one correlation between abilities not very like one another which could not be explained as due to the general factor.

In 1912, Spearman and Hart (**3**) summarized the articles on this subject which had appeared previously and calculated the intercolumnar correlation coefficients for several tables of correlations in these earlier studies. In this paper, they presented as the fundamental criterion of hierarchy the equation,

$$\frac{r_{ac}}{r_{ad}} = \frac{r_{bc}}{r_{bd}}, \tag{1}$$

in which *a*, *b*, *c*, and *d* are four variables or tests. By the use of partial correlation formulæ, they furnished a simple proof. The proof was of the following type. If each variable of a group of variables is composed of a general factor, '*g*,' and an uncorrelated specific factor, '*s*,' then the partial correlation between any two of these variables with '*g*' eliminated is zero. The only way in which the expression

$$r_{ac \cdot g} = \frac{r_{ac} - r_{ag}r_{cg}}{k_{ag}k_{cg}} = 0 \tag{2}$$

may be true is that the numerator be equal to zero or that

$$r_{ac} = r_{ag}r_{cg}. \tag{2a}$$

Similar expressions can be formed from any combination of two of these variables. If between such expressions as (2a), one forms the ratios,

$$\frac{r_{ac}}{r_{ad}} = \frac{r_{ag}r_{cg}}{r_{ag}r_{dg}} \quad \text{and} \quad \frac{r_{bc}}{r_{bd}} = \frac{r_{bg}r_{cg}}{r_{bg}r_{dg}}, \tag{2b}$$

then (1) follows at once. The converse proposition that, if (1) is true, 'g' and 's' alone determine the observed variables was assumed in this article and not proved until later (**12, 18, 23**).

Spearman and Hart used the intercolumnar correlation coefficient as an approximation to the fundamental criterion of hierarchy (1) because they possessed at this time an expression for the probable error of the former but not for the probable error of the latter. The theory of a general factor demands that, omitting pairs containing r_{pp},

$$Rr_{ap}r_{bp} = 1 \quad (p = a, b, \cdots, n). \tag{3}$$

When (3) tends to be true, the intercorrelations are said to be hierarchical. When the more exact criterion (1) is true, the intercorrelations are said to be equiproportional. Hierarchy is the more general and looser term. It will be shown later, that if (1) is true, (3) is always true (unless it is indeterminate), but the converse is not necessarily true.

It was shown that R must be zero if the non-focal theory of intelligence is to hold; that it must be negative for the multi-focal theory of non-overlapping factors (Thorndike's early view); and it must be unity for the unifocal theory of a general factor with no group factors (Spearman's view). The case of overlapping factors in a multi-focal theory was not treated here but was developed later by Thomson as his Sampling Theory (**10**).

On applying this more precise method of inter-columnar correlation to all the tables of intercorrelations of mental and sensory traits that had been published up to that time, a striking unanimity was revealed. Previously, the degree of hierarchy had varied with the subjective estimation of the inspector and his arbitrary arrangement of the columns.

The average inter-columnar correlation from the tables of fourteen different investigators, summarizing thirty years of psychological researches and representing a great wealth of test material, was unity, as expected by the unifocal hypothesis of a general factor. It seemed to be the most striking quantitative fact in the history of psychology.

It must be noted that this finding was impaired by two procedures. The first was that the tests 'too much alike' were pooled, or one omitted; for, in them, the specific factors overlapped and so produced additional correlation. (We may designate these group factors *o*, analogous to the '*g*' and '*s*' factors.) The other procedure open to criticism was that some of the columns were rejected as falling below the 'correctional standard.' Some such standard was required because the inter-columnar correlation between any two columns tends to become indeterminate as the correlation coefficients within a column tend to become equal. The standard fixed upon was to use only columns in which the probable error of the correlations about the mean correlation of the column was twice the mean probable error of the correlations in the column.

Spearman and Hart then pointed out that the Binet tests owed their success to the real abandonment of Binet's expressed multi-focal or faculty theory of intelligence and to the implicit adoption of the unifocal theory. By using a great diversity of tests, the independent specific factors neutralized each other while the general factor became more prominent. They pointed out that '*g*' has much in common with 'clear awareness' and 'attention,' though it is much more than this. Its amount appears to be quite constant within the individual. It enters into the higher or central mental functions to a far larger extent than into the peripheral or sensory functions. They interpreted it in psychological terms as 'intellective energy.' They attributed, in part at any rate, the variations of this energy to variations in the bloodstream with its nutritive and glandular constituents. They stated that "physiology, no less than psychology, indicates the possibility of the General Factor having a more or less composite origin" (p. 72).

In appendices, they furnished mathematical proofs of the inter-columnar correlation criterion, the correction for attenuation and formulæ for probable errors of sampling. These will be further discussed when, in following the chronological order of this review, the criticisms of these proofs are reviewed.

Two years later, Spearman (**5**) applied the inter-columnar correlation criterion to Simpson's data. The latter, working under Thorndike's direction, had asserted that no hierarchy existed in his correlations. But, when the mathematical criterion was used instead of mere inspection, the inter-columnar correlation proved to be almost unity, and the hierarchy therefore very nearly perfect. This article repeated the proof of (1), gave the formula for the inter-columnar correlation criterion corrected for attenuation, and an example of the arithmetic of its detailed application. Spearman also brought forward in this article (**5**) "the not implausible hypothesis that each performance depends on a randomly selected group of very numerous independent elements, and that the correlation between any two performances is due to some of the elements happening to be common to both groups" (p. 109). But this hypothesis is rejected on the grounds that by it

$$Rr_{ap}r_{bp} = r_{ab}.$$

In another contribution (**6**), Spearman pointed out that discussions in the field of eugenics were handicapped by the divergent hypotheses concerning abilities. There were: (*a*) the hypothesis of compensation, in which an individual inferior in some traits tended to be superior in others; (*b*) the hypothesis of independence of abilities (uncorrelated or non-focal); (*c*) the hypothesis of faculties (multi-focal or non-overlapping); (*d*) the hypothesis of a general factor (unifocal theory). He then reviewed the evidence concerning the 'two factor theory' which combines hypotheses (*b*) and (*d*).

The equation for an individual's amount of the general factor, '*g*,' was given as

$$g_x = a_x r_{ag} \pm .6745\sigma_a\sqrt{1 - r^2_{ag}}, \qquad (4)$$

where g_x is measured from the mean of all individuals and is the amount of 'g' possessed by the individual x, with an unreliability as indicated by its probable error of estimate.

By combining three equations such as (2a), r_{ag} may be found from the expression

$$r^2_{ag} = \frac{r_{ab}r_{ac}}{r_{bc}}, \tag{5}$$

where a, b, and c are any three tests of a larger number of tests which give correlation coefficients satisfying the criterion of equiproportionality. Although not stated by Spearman until later, it may be noted here that the basic equation defining the theory of 'g' is allied to (4), namely:

$$a_x = r_{ag}g_x + r_{as}s_{ax} = r_{ag}g_x + \sqrt{1 - r^2_{ag}}s_{ax}, \tag{6}$$

where in addition to the symbols explained above, s_{ax}, is the specific factor in a particular test, a, possessed by an individual, x, and all scores are in standard deviation units. The independence of 'g' and 's' is here implicitly expressed, for otherwise (6) could not be a true equation.

Spearman then discussed the influence of environment on 'g.' He stated that the amount of an individual's 'g' is, for the most part, determined by heredity, but that the 's' is more influenced by nurture. He also advanced the hypothesis that 'g' increases from birth until puberty. Though recognizing the faulty nature of the experimental evidence, he argued that if 'g' were partly due to nurture its effect would vary in different tests in which the individual's training differs. Since 'g' is, by definition, that ability of an individual which is constant from test to test (though each test may be dependent upon the 'g' factor to different degrees indicated by r_{ag}), that which varies from test to test—such as effects of differential training—would be an 's' and not a 'g.'

In 1916, the theory of 'g' was effectively challenged by Thomson in a paper (7) dealing with correlations derived from a system of dice. He agreed that, if 'g' and 's' existed solely, the observed intercorrelations would be in perfect

hierarchy, but he argued that hierarchy could also be produced by group factors without a general factor.

He used his dice in groups and derived the formula[2]

$$r_{xy} = \frac{c}{\sqrt{(c+m)(c+n)}}, \tag{7}$$

where c is the number of dice common to the two variables x and y, and m and n are the number of residual dice in x and y respectively. An x score is the sum of the dots on the faces of the dice turned up in the throw for one individual's x score; a y score is similarly constituted. The common dice, c, are tossed only once for each individual, hence the common dice enter into both the x and y scores with identical values. The relationships introduced by using dice will be pointed out in the summary of properties of factors at the conclusion of this review. With 145 dice, of which 36 were group factors (visualized by combinations of colors of dice), he artificially produced 10 symmetrically distributed variables corresponding to 10 tests. He used 36 actual tosses of the dice corresponding to scores from 36 individuals. No dice were common to all the variables so that there was no general factor. By the theoretical correlation formula (7) and by actual tosses of the dice, he found that the intercorrelations of the variables, both raw and corrected for attenuation, yielded inter-columnar correlations approximating unity.

In a brief note (8), Spearman objected that these dice arrangements were arbitrary, undefined, and extremely improbable by chance. Thomson agreed that the dice arrangements were artificial, but he argued that the important point was the production of a hierarchy, by means other than a general factor.

Next, Thomson brought forward more evidence (10) that hierarchies could be produced by group dice factors without a general dice factor. (Note that the word 'dice' is here used by the reviewer as it should have been used in every case in

[2] Spearman has pointed out that this formula was expressed by him in theorem form in 1904—that a correlation was the geometric mean of two percentages of common elements—and that Thomson's formula follows directly from his formula for the correlation of sums.

Thomson's discussions, for, a general factor of the '*g*' type was later proved to exist (17) even if no general dice factor was present.) In reply to Spearman's objection of 1916 that his hierarchy from dice represented extremely improbable situations, Thomson showed that they also could be obtained on chance expectations. But Thomson's arrangement of chance influences was very different from Spearman's, and this makes an enormous difference in the results. Thomson secured excellent hierarchies by letting chance, operating as outlined below, determine the amount and overlap of the dice factors.

The procedure was first to determine the number of group dice factors and the number of specific dice factors in each variable by draws of cards from a constant pack. Thus, for variable *I*, first a 5 and then an 8 might be drawn. This would mean that variable *I* would have five group or overlapping factors and eight specific ones. The next step was the determination of the group factors. To do this, one suit was used; it was well shuffled, and from it there were withdrawn, for any particular variable, cards equal in number to the number of group factors which had previously been determined for this variable. These cards were noted down. They were then replaced and shuffled, and a new drawing was made for another variable. The next step was the counting of the number of factors common to any two variables as the determination by formula (7) of the correlation between them. Such chance determined variables[3] contained no general dice factor, yet they gave intercorrelations the intercolumnar correlations of which approached unity fully as well as did the hierarchies from psychological data summarized by Spearman and Hart in 1912. The reconciliation of the

[3] The illustration would have been much improved by not drawing for the number of specific factors but letting each variable be determined solely by the draws for its group. Professor Thomson writes to the reviewer saying, "I quite agree that it would have been better had I not drawn for specific factors: and that it would have been better not to have spoken of two levels, but, as I later did, to allow any factor freedom, when naturally some would have been specific and all would vary in what you call 'order' or 'width.' See *Brit. J. Psychol.*, July 1920, 10, 325 and May 1919, 9, 341 near middle of page. 'Indeed the specific factors may be entirely omitted. . . .'"

'*g*' theory with this apparent contradiction will be outlined in reviewing Garnett's work of 1919.

On this evidence, Thomson proposed a sampling theory of ability to explain the observed facts of hierarchy. By the sampling theory, the mind is supposed to have two levels. One is the upper, central level in which the elements, which are presumably cortical neurones, behave as group factors, that is, as elements entering by means of various groupings into more than one type of measured behavior. The other is the lower, more peripheral, and sensory level in which the elements tend more to be specific to each variable. Neurones are suggested as the elements, as they function in an all-or-none manner, thus bearing a close analogy to the dice and cards referred to before. By this sampling theory, a form of behavior is determined by a sample (or 'team' or 'pattern') of neurones. Correlation is due to overlapping samples of neurones; transfer of training and improvement due to practice are a result of a 'shaking down' to a more economical sample. Instead of two forms of factors, there may be any number of them, varying in scope from the specific to the general.

Thomson argued that since the facts of hierarchy are compatible either with Spearman's view of general and specific factors or his own view of group dice-like factors, then, either 'intellective energy' or 'samples of neurones' could be the interpretative hypothesis for psychology. He felt that the view of neurone samples had the better physiological basis.

In an article in *Biometrika* (**16**), Thomson attacked the inter-columnar correlation as a criterion of hierarchy. His position was that the 'correctional standard' accepted only columns which naturally tended to give correlations averaging around unity, for, the distribution of the corrected values extends from zero to infinity. His diagram is reproduced in Fig. 1. In practice, the 'correctional standard' tended to limit the corrective values to numbers within equal distances on either side of unity, thus leaving a distribution with one as an average.

In Appendix A of Spearman and Hart's 1912 article, the quantity ($\Sigma\rho'\epsilon$), the sum of the products of the correlation coefficients and their errors, was treated as negligible on the assumption of random errors which would therefore be un-

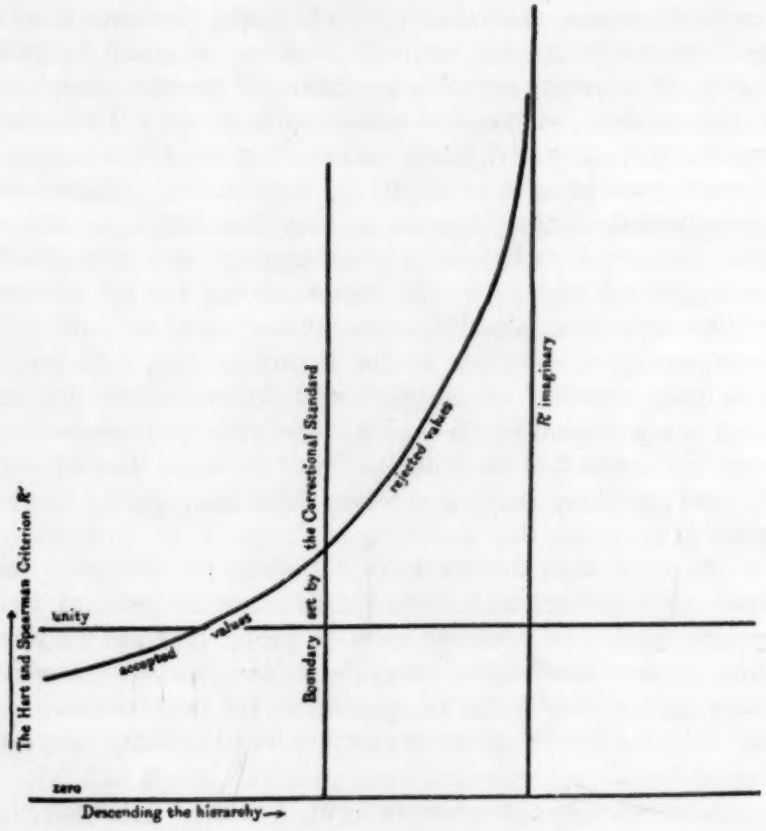

FIG. 1.

correlated with the ρ's. Thomson argued that these errors are correlated and tend to be of one sign. He gave an example from dice in which an average inter-columnar R_{rr} of .59 had been corrected to an R_{rr} of 1.00 because the correctional standard had rejected the columns with inter-correlations far below unity; thus a bad hierarchy had been changed into an apparently good one.

The application of the formula for the correction for attenuation to the inter-columnar correlation was criticized by several people. Pearson pointed out that a correlation above unity was impossible, even within probable error limits, inasmuch as the probable error approaches zero as the correlation approaches unity. In 1912, Spearman and Hart averaged correlations some of which were greater than unity and thus obtained a correlation equal to one. This use of theoretically impossible correlation coefficients cast further doubt upon the efficacy of the inter-columnar correlations as a criterion of hierarchy. In support of Spearman, it might be added that if the formula for correction gives a correlation above unity, the formula is not necessarily incorrect, but the conditions under which it was developed are not fulfilled in the given data. Spearman argued that we need not call this index a correlation coefficient but simply a function the average of which approximates the true correlation coefficient.

Thomson suggested (11) that hierarchy is produced by errors of sampling the elements in correlations. The correlation coefficients are themselves correlated, to a degree shown by Filon and Pearson's formula. In a brief explanation of this hypothesis, we may start with a given number of tests the correlations of which, when calculated from an infinite sample, would be equal. In any finite sample, however, one correlation coefficient, let us say r_{38}, will be largest simply because of the errors involved in the sampling of the individuals. This fact will tend to make all the correlations with 3 and 8 erroneously high, thus making 'ridges' along the 3 and 8 arrays with a 'node' at r_{38}. The next higher coefficients would also tend to form ridges and nodes in the order of their observed magnitudes, and thus hierarchy would be shown. When this reasoning is applied to the case in which the correlations from an infinite sample are not equal, the same tendency to form ridges and nodes —'steeper' hierarchies—would be observed in finite samples.

Thomson called attention to two difficulties in his proposal: (*a*) Filon and Pearson's theorem of the correlation of correlation coefficients was developed for correlations arising

from populations distributed normally and not for correlations derived from dice-like elements; and (*b*) in sampling a finite population all the correlations should tend to rise or to fall, as a whole, but hierarchy is produced by some correlations doing one thing and some the other. To the reviewer, it seems highly unlikely that hierarchies, of the perfection and steepness often observed, can be due entirely to random errors of sampling the elements (*e*'s) though this situation would exaggerate the true values. There is required some bias, some mechanism, for making this sample not purely random. This, Thomson suggests, is in our choice of tests which is determined by many considerations not at all of a random nature. The reviewer fully agrees with this.

It may here be stated, parenthetically, that hierarchy is often visualized as the 'sloping off' tendency in a table of intercorrelations. That is, there is a tendency for the correlations to become smaller from left to right and from top to bottom when the largest coefficient has been placed in the upper left hand corner. From a table of equiproportional correlations, there may be derived for each array or test a coefficient unique to that array. The product of two of these coefficients will be the correlation between these two tests or, in other words, the number at the intersection of these two arrays. In each case, this *array coefficient*, as it may be called, is identical with r_{ag} (5), the correlation between a test and a general factor when (1) is true. The proper arrangement of a table of correlations is in order of the size of these array coefficients.[4] In a table of equiproportional coefficients, the $n(n-1)/2$ correlations are completely specified by the n array coefficients. There are many uses and interpretations of this array coefficient of which one of the most important is its use as a measure of the 'sloping off' tendency of the table.

Thomson investigated the conditions (9) which permit one to use the intercorrelations in inferring a general dice factor from three variables. He used the following illus-

[4] The method of their determination is described in Appendix IV of 'The Abilities of Man,' Spearman.

tration to show how ambiguities might arise from the analysis of variables into factors when partial correlation coefficients are used.

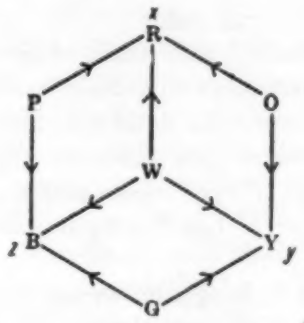

FIG. 2.

The letters *R*, *O*, *Y*, *G*, *B*, *P*, and *W* represent groups of dice (red, orange, yellow, green, blue, purple and white). Variable *x* is the sum of the *P*, *R*, *O*, and *W* dice; *y* of the *O*, *Y*, *G*, and *W* dice; and *z* of the *G*, *B*, *P*, and *W* dice. *W* is thus a general dice factor, *O*, *G*, and *P* are group dice factors, and *R*, *Y*, and *B* are specific dice factors. For simplicity of illustration, let the seven dice factors contain the same number of individual dice. Now, if the group factors *O*, *G*, and *P* are deleted, the three variables *x*, *y*, and *z* will show intercorrelations of .50, which are due to the general factor, *W*. But alternatively, by deleting the specifics, *R*, *Y*, and *B* and the general dice factor *W*, there result correlations of .50 as before, but this time due solely to the group dice factors, *O*, *G*, and *P*. The partial correlation between any two of the dependent variables with the third controlled is .33 which in no sense affects the elimination of the factors common to the variable controlled and the other two.

In such dice situations, an infinity of arrangements can be made, each of which leaves the correlations and the total number of dice in each dependent variable unchanged. The determination of variables by dice-like elements and factors is not unique but ambiguous. This is always true of dice factors in any number of variables. It is also true of the

'g' and 's' types of factors if less than four variables are used. With four or more variables, and the tetrad equation (1) equal to zero, the 'g' and 's' as defined by Garnett (12) become uniquely determined.

Thomson concluded that from observed correlations we cannot infer the structure of factors uniquely. This is true of dice factors because the determination of the correlations by different structures give different values for the factors. However, if 'g' and 's' are present and if there are no residual group factors after 'g' has been partialled out, the 'g' and 's' are unique.

To this same article (9), he wrote, with J. R. Thompson, an appendix in which was developed a criterion for the necessary presence of a general dice factor. This criterion was deduced from the equation which expresses the condition that the correlation be due only to group dice factors (such as O, G, and P in Fig. 2). This criterion is

$$r^2_{12} + r^2_{13} + r^2_{23} + 2r_{12}r_{13}r_{23} - 1 = O = F. \qquad (8)$$

Equation (8) is the expansion of the determinant expressing the conditions that three equations similar to (7) be true if the correlations are due only to group factors. F can become greater than O only by having the correlations increased; this can occur only through the introduction of some general, W, dice factor into the situation. Likewise, F can become negative only by the introduction of specific dice factors R, Y, or B, to lower the intercorrelations. Thus, if F is positive, a general dice factor is necessarily present, and if F is negative, specific dice factors are necessarily present. In any case, however, additional general or specific dice factors may be present if offset appropriately by the other. We may note in passing that Thomson's criterion is not limited to dice factors solely.

The development of the theory of 'g,' 'o,' and 's' [5] was

[5] This designation is convenient because it allows the combination of Spearman's theory of two factors, 'g' and 's,' with the theory of overlapping factors, 'o.' The elements in Thomson's sampling theory may be designated as 'd' factors. This combination theory is more general than the existence of hierarchy. The group factors similar to 'o' and the group factors similar to 'd' will be distinguished from

greatly advanced at this point by Garnett. He started (12) with the same assumptions as those used by Bravais in 1846 in deriving the normal correlation surface, namely: that a normally distributed variable can be expressed as a function of many small variables distributed normally. Taylor's theorem is assumed applicable whereby an additive combination is used as a sufficiently satisfactory approximation for whatever the real functional combination may be. The n-fold correlation surface is then derived and the probability of any point or 'density' of points is formulated.

It had previously been proved that hierarchy existed if 'g' and 's' alone determined the intercorrelations, but the converse had been assumed. In this article, Garnett proved the converse theorem for normally distributed variables. He established the proposition that if equiproportion exists, only 'g' and 's' are present, and there are no 'o' factors. Furthermore, he rigorously proved that 'g' and 's' are uniquely determined and additively combined. If, instead of the tetrad difference, the inter-columnar correlation be used—as is rarely done now—a complicating $(n + 1)$th variable entered into the proof. This last variable however also expressed the exclusive and unique existence of 'g' and 's,' but it expressed them as factors in which the n correlated variables might be reëxpressed.

In this article (13), Garnett again derived the 'cosine law' and the fundamental system of equations for the theory of direction cosines in n-dimensional space. These constitute the geometric interpretation of correlation. Thus, any normal variable, q, can be expressed as a vector by N coördinates mutually at right angles in n-space.

$$q_1 = r_{e_1 q_1} e_1 + r_{e_2 q_1} e_2 + r_{e_3 q_1} e_3 + \cdots + r_{e_N q_1} e_N. \qquad (9)$$

The r's are the cosines of the angles between the e coördinates and the vector q_1. All the e's [6] are uncorrelated because the

each other in the concluding part of this review. Note here, however, that the former are group factors remaining after 'g' has been eliminated and that the latter are group factors which are transformations of 'g,' 'o,' and 's.'

[6] The e's denote independent elements of which a die, d, is a particular variety having six equiprobable values and a rectangular distribution.

cosine of a right angle is zero. In such an orthogonal system, it is always true that

$$r^2_{e_1q_1} + r^2_{e_2q_1} + r^2_{e_3q_1} + \cdots + r^2_{eNq_1} = 1. \qquad (10)$$

Each coördinate or independent factor, e, contributes a percentage amount, r^2_{eq}, to the determination of the dependent variable, q.[7]

Note that this r_{eq} or cosine, r_{aq} of (6), and the array coefficient from an equiproportional table are all the same thing. Furthermore, r^2_{eq} has all the advantages of being a percentage, which advantage correlation coefficients do not have. This is due to the independence of the e's and the complete dependence of q upon them. The equation (6), which defines the theory of 'g,' is but a special case of (9) in which $N = 2$, $g = e_1$, and $s = e_2$. When the correlations of the q's are equiproportional, the equations for all of them will have e_1 unchanged, but the e_2 of (9) or the 's' of (6) and coefficients r_{eq} will change from one q to another and from e_1 to e_2. We have, then, in (9) the general equation by which one can postulate and test various hypotheses concerning the combinations and arrangements of factors of the 'g-o-s' type. There is here a tool of possibly immense value for the quantitative analysis of psychological data.

Garnett's 'cosine law' states the correlation between any two observed variables, q_1 and q_2, as

$$r_{q_1q_2} = r_{e_1q_1}r_{e_1q_2} + r_{e_2q_1}r_{e_2q_2} + \cdots + r_{eNq_1}r_{eNq_2}. \qquad (11)$$

Mathematically, this is the cosine of the angle between two vectors. It can be derived from Spearman's formula for the correlation of sums of variables; it is the numerator of that formula when the e's are uncorrelated and the sigmas of the composite are taken as units. It is the general case of (2a), in which the r's involving the independent 's' factors become zero leaving only the first term on the right of (11).

To test hypotheses of factorial structure, we have only to test the r_{eq}'s in such equations as (9) in any desired way. Thus, if r_{e_1q}, for all the q variables, never equal zero, then

[7] This theorem was stated by Spearman in 1904.

e_1 is a general factor. If it equals zero in every q except one, it becomes a factor specific to that q. Similarly group factors may be discovered. The order of a group factor would be designated by the number of q's into which it entered, and its size would be indicated by its r_{eq}. There should be derived, as far as possible, criteria for the various relationships which would permit inference as to the structure of the factors. There are, however, definite limits to this procedure if the number of unknown r_{eq}'s exceeds the number of known r_{qq}'s.

As a corollary to (11), Garnett discovered another useful law—when three variables represented by three vectors happen to lie in one plane, they can be perfectly expressed in terms of two general factors, two coördinates of that plane. The necessary and sufficient condition is that the three angles between these vectors (taking the largest angle negatively) shall add to zero. Consider the three observed correlations as cosines and from a table of natural cosines convert them into angles. Then, by inspection of the sizes of these angles, there can be determined which sets of three variables can be factored into two general factors. In terms of angles, this criterion is:

$$\text{Angle } A + \text{Angle } B - \text{Angle } C = 0 \qquad (C \text{ being the largest}) \quad (12)$$

In terms of correlation coefficients, this is written:[8]

$$1 - r^2_{12} - r^2_{13} - r^2_{23} + 2r_{12}r_{13}r_{23} = 0. \qquad (13)$$

In a footnote, Garnett derived from the formulæ of spherical trigonometry the critical equation for factoring four variables into three general factors. With more than three variables this type of criterion cannot be expressed in simple terms of angles as in (12), but it can be expressed in terms of correlation coefficients analogous to (13).

In Section V, Garnett applied criterion (12) to Webb's study, which provided the intercorrelations of forty-eight

[8] Note the difference in sign which distinguishes (13) from (8). These are very different criteria though similar equations; both are expansions of third order determinants.

traits, measured by subjective but exhaustive ratings of college men by their acquaintances over months of observation and study. He found four traits, designated quickness of observation, originality, sense of humor, and '*g*' (from intelligence tests) which, when taken three at a time, approximately satisfied (12) and could be expressed by means of two general factors. The factor '*g*' was chosen as the first general factor, and its perpendicular coördinate was determined from the correlations and designated cleverness, '*c*.' The correlations of the variables with *c* are the cosines of the angles between the vectors of the variables and the axis of '*c*.' They are identical with a first order regression coefficient of '*c*' on *x* with '*g*' controlled.

Webb's 'purpose' or 'persistence of motives' factor is shown by Garnett to be (with a slight alteration to make it independent of '*g*' and '*s*') another independent factor common to a large number of character traits and which, with '*g*' and '*c*,' completely specify many traits.

As soon as sufficiently reliable data are available, this method of analysis will offer a means of expressing the chaotic mass of the present complex traits in fewer and non-overlapping concepts. In time, it may become possible to build a systematic psychology upon mathematical equations, but probably there will never be as close an agreement between the theory and the observed data as could be desired.

In a third contribution, (17), Garnett showed the relations existing between the factors '*g*' and '*s*' and the factors similar to dice factors. He demonstrated that both types of factors were consistent with the facts of equiproportion. Either set of factors was a mathematical transformation of the other.

In this article, he demonstrated that n correlated variables may be expressed in terms of n independent variables which may be selected with $n(n-1)/2$ degrees of freedom. Geometrically this is the conversion of n vectors into their components along n mutually perpendicular coördinates. He next demonstrated that there are $n(n-3)/2$ independent conditions which must be satisfied if equiproportion is to

exist throughout the table of intercorrelations. For less than four variables, there are no conditions which need to be fulfilled; thus, there are an unlimited number of ways in which three variables may be separated into '*g*' and '*s*.' To obtain a unique '*g*' and '*s*,' there must be at least four variables.

He showed that '*g*' and '*s*' could be obscured and that group dice factors, as in Thomson's experiments, be made to appear by analyzing the variables, which had given equiproportional correlations into a large number, n, of equal variables, e, such as dice elements, instead of analyzing the variables into a '*g*' and an '*s*.' Since the elements of (9) are equal, the r_{eq}'s are equal, and by (10)

$$r^2_{eq_s} = \frac{1}{N} = \frac{1}{np_s} \qquad \text{or} \qquad r_{eq_s} = \frac{1}{\sqrt{N}} = \frac{1}{\sqrt{np_s}}. \tag{14}$$

In this equation, N is the number of dice in the variable, q_s, and we may think of this N as being a proportion, p_s, of the total number, n, of dice available; there will, of course, be $(1 - p_s)n$ dice not entering into the variable, q_s. Equation (9) may now be written

$$q_s = \frac{1}{\sqrt{np_s}}(\text{sum of } np_s \text{ whole } e\text{'s}) \quad (s = 1, 2, \cdots, n). \tag{15}$$

The number, n, must be very large, approaching infinity unless further assumptions are made regarding the weights of the e's and the determination of p_s. Equation (15) is the fundamental equation for defining the sampling theory, just as (6) is the fundamental equation for the theory of '*g*.' Each is a particular form of the general equation (9).

In this same article (17), Garnett showed that the most likely number of elements common to the two variables, q_s and q_t, would be np_sp_t. The correlation between q_s and q_t follows at once

$$r_{q_sq_t} = \frac{np_sp_t}{\sqrt{np_s}\sqrt{np_t}} = \sqrt{p_sp_t}. \tag{16}$$

Thus, if equiproportion exists, the correlation coefficient is the

square root of the probability of there arising, by chance, elements common to the two variables in question. In this sentence, the usual laws of probability are recognized.

It follows at once from

$$\frac{r_{st}}{r_{su}} = \frac{\sqrt{p_s p_t}}{\sqrt{p_s p_u}} = \frac{r_{tv}}{r_{uv}} = \frac{\sqrt{p_t p_v}}{\sqrt{p_u p_v}}, \qquad (17)\ ^{9}$$

that the most probable value of the tetrad difference is zero for variables correlated in a chance manner and dependent upon numerous dice-like elements.

Garnett gave an equation [10] and a Rule, showing in detail what might be expected as to the number of factors specific to one test, common to only two tests, common to only three tests, and likewise for all degrees of commonality. By this expression, there can be produced from dice-like factors a table of perfectly equiproportional correlation coefficients without trial and error methods and even without using the theoretical correlation formulæ such as (7). It should be noted that the number of general dice factors is $np_1p_2\cdots p_n$; this number, compared to the residual number of group dice factors, is very small, because the p's are proper fractions. In the approximations of practice, this general dice factor may disappear as it did in Thomson's experiments and still satisfactory hierarchies may result.

In order that absolutely perfect equiproportion exist, the number, n, of dice factors may have to be very large. The required number is the least common multiple of the p's for all the tests, but in practice 100 or even a smaller number frequently give results indistinguishable from the results obtained from psychological data. Until r's can be reliably calculated to six or seven decimal places, a finite number of dice-like elements is often adequate.

In Section V, Garnett gave the exact equations [11] trans-

[9] Note that $\sqrt{p_s} = r_{sq}$. The symbol, s, here stands for the variable q_s and not for the specific factor, 's.'

[10] For the c_s in his notation, we have substituted p_s in order to connect this quantity with the common concepts of percentage and probability. The n of this article is the same as the N of his.

[11] Three misprints may puzzle anyone who checks Garnett's proof between his equations (17) and (18) on p. 251. These do not change the validity of (18) which is correct as printed.

forming the dice-like e's into 'g' and 's' and vice versa. These are, for the variable, q_a,

$$g = \frac{\Sigma(e)}{\sqrt{n}} \qquad \text{or} \qquad g = \frac{\Sigma(d)}{\sqrt{n}}, \tag{18}$$

$$s_a = \left[\frac{\sqrt{1 - p_a}}{\sqrt{np_a}}(\text{sum of the } np_a\ e\text{'s in } q_a)\right] - \left[\frac{\sqrt{p_a}}{\sqrt{n(1 - p_a)}}(\text{sum of the } n(1 - p_a) \text{ remaining } e\text{'s})\right] \tag{19}$$

$$(a = 1, 2, \cdots, N).$$

Since the 'g' and the e's are interchangeable, it is unnecessary for mathematics that a particular e appear in every variable, but it is necessary that the e's be distributed in accordance with the usual laws of probability. This freedom is well illustrated in Thomson's dice structures. For neurones, there would be more restrictions.

Thus, 'g' is a measure of the proportion of shared elements. It is a concept expressing very simply the quantitative effect which may be caused by numerous all-or-none elements such as neurones. In Section VII, Garnett ventured a possible interpretation of 'g' in psychological terms: if 'g' "measures the subject's Will or power to concentrate in a particular system of nervous arcs, the various e's may measure the numbers of a subject's neurones rendered active, according to Lucas and Adrian's 'all-or-none' law, in that subject's exercise of his different qualities" [12] (p. 256).

In sum, this third contribution of Garnett stated that mathematically, the facts of equiproportion could be stated either in terms of 'g' and 's' factors or in terms of e factors distributed according to the laws of probability. He preferred to use 'g' and 's' because they were simpler than the e's. It does not follow, however, that in physiological or psychological terms the two interpretations would be the same.

Spearman noted that Garnett had not assumed a complete 'chance' system. Garnett had required, for the system of

[12] The reviewer has replaced Garnett's 'x's' by 'e's' for uniformity of symbolism.

equations of the form of equation (9), that the coefficients of the *e*'s be uncorrelated when the variables are taken in pairs of rows. He had not required that these coefficients of the *e*'s be uncorrelated when taken in pairs of columns. It seems likely, however, that if he had done so, he would have secured the same results. This may be inferred from Garnett's more general work and also from the fact that Thomson's dice structures satisfied the tetrad difference criterion.

In reply (**20**) to this analysis, Thomson argued: (*a*) that the inter-columnar correlation was an unsatisfactory criterion of hierarchy; (*b*) that Garnett's elaborate Rule and conditions for all-or-none elements to form equiproportional correlations were simply the laws of probability; and (*c*) that at least in one case even the all-or-none provisions were not essential as unequally weighted or fractional dice factors gave the same results. Thomson's view was that the '*g*' and sampling theories were not alternative modes of expression but were species and genus—squares and quadrilaterals—the '*g*' theory being the special and limiting case when the dice factors were very numerous and approached pure probability in arrangement.

Thomson further pointed out that Garnett's linear transformations (18) and (19) were completely successful only if perfect equiproportion existed. This is never directly observed because of sampling errors. Finally, he objected that it was much less real to deal with a transformation, a '*g*' which corresponded to "a little bit of this neurone, and a little bit of that" and not necessarily to the actual determiners. To explain that which enables one individual's nervous system to build up its reaction patterns to greater complexity than another individual, he preferred the more concrete case of *e* elements to the more abstract case of '*g*' and '*s*' factors. He could then think of the causes as being neurones, Mendelian units, or some similar entity.

In 1919, Garnett and Thomson wrote a joint note (**14**) to show that they were really far more in agreement than seemed possible from their separate articles which had been

devoted to developing points of difference. They agreed on the general equations (9) and (10) with (15) as the special case of *d* elements expressing the sampling theory. They agreed that equiproportion could be expressed, mathematically, either by '*g*' and '*s*' factors or by dice factors, which are distributed according to chance and theoretically infinite in number. As equiproportion becomes perfect, the *d* elements tend to become constituent elements or units of '*g*.' The two authors differed in the psychological interpretations which they preferred. Although Thomson agreed to the statement concerning the transformations, he objected to the interpretation of it. The dice in the form in which they exist physically are not elements of '*g*,' but they can be transformed mathematically into it. They are as bricks in a wall; each brick may be a foot long and can be transformed into meters. The length of the wall can then be given in meters, but it is more naturally given in feet because the wall is built of foot bricks. This analogy shows a difference which is largely one of arithmetic labor. Thomson feels, however, that in the case of the '*g*' and *d* elements we are dealing with new and little explored entities, and the form of expression may make large differences in interpretation and further experiment.

REFERENCES

1. SPEARMAN, C., '"General Intelligence," objectively Measured and Determined,' *Amer. J. Psychol.*, 1904, 15, 201–293.
2. KRUEGER, F., AND SPEARMAN, C., 'Die Korrelation zwischen verschiedenen geistigen Leistungsfähigkeiten, *Zsch. f. Psychol.*, 1906, 44, 50–114.
3. HART, B., AND SPEARMAN, C., 'General Ability, its Existence and Nature,' *Brit. J. Psychol.*, 1912, 5, 51–84.
4. HART, B., AND SPEARMAN, C., 'Mental Tests of Dementia,' *J. Abn. Psychol.*, 1914, 9, 217–264.
5. SPEARMAN, C., 'The Theory of Two Factors,' PSYCHOL. REV., 1914, 21, 101–115.
6. SPEARMAN, C., 'Heredity of Abilities,' *Eug. Rev.*, 1914, 6, 219–237.
7. THOMSON, G. H., 'A Hierarchy without a General Factor,' *Brit. J. Psychol.*, 1916, 8, 271–281.
8. SPEARMAN, C., 'Some Comments on Mr. Thomson's Paper,' *Brit. J. Psychol.*, 1916, 8, 282–284.
9. THOMSON, G. H., 'The Proof or Disproof of the Existence of General Ability,' *Brit. J. Psychol.*, 1919, 9, 321–336.
10. THOMSON, G. H., 'The Hierarchy of Abilities,' *Brit. J. Psychol.*, 1919, 9, 337–344.
11. THOMSON, G. H., 'On the Cause of Hierarchical Order among the Correlation Coefficients of a Number of Variates taken in Pairs,' *Proc. Roy. Soc. Lond.*, 1919, 95A, 400–408.

12. GARNETT, J. C. MAXWELL, 'On Certain Independent Factors in Mental Measurements,' *Proc. Roy. Soc. London*, 1919, **96A**, 91–111.
13. GARNETT, J. C. MAXWELL, 'General Ability, Cleverness and Purpose,' *Brit. J. Psychol.*, 1919, **9**, 345–366.
14. GARNETT, J. C. MAXWELL AND THOMSON, GODFREY H., 'A Joint Note on "The Hierarchy of Abilities,"' *Brit. J. Psychol.*, 1919, **9**, 367–368.
15. THOMPSON, J. RODLEY, 'The Rôle of Interference Factors in producing Correlation,' *Brit. J. Psychol.*, 1919, **10**, 81–109.
16. THOMSON, GODFREY H., 'On the Degree of Perfection of Hierarchical Order among Correlation Coefficients,' *Biometrika*, 1919, **12**, 355–366.
17. GARNETT, J. C. MAXWELL, 'The Single General Factor in Dissimilar Mental Measurements,' *Brit. J. Psychol.*, 1920, **10**, 242–258.
18. SPEARMAN, C., 'Manifold Sub-theories of "The Two Factors,"' PSYCHOL. REV., 1920, **27**, 159–172.
19. THOMSON, G. H., 'General versus Group Factors in Mental Activities,' PSYCHOL. REV., 1920, **27**, 173–190.
20. THOMSON, GODFREY H., 'The General Factor Fallacy in Psychology,' *Brit. J. Psychol.*, 1920, **10**, 319–326.
21. YULE, G. UDNY, 'Critical Notice of "The Essentials of Mental Measurement,"' *Brit. J. Psychol.*, 1921, **12**, 100–107.
22. GARNETT, J. C. MAXWELL, 'Review of "The Essentials of Mental Measurement" by William Brown and Godfrey H. Thomson,' *Brain*, 1921, **44**, 332–337.
23. SPEARMAN, C., 'Correlation between Arrays in a Table of Correlations,' *Proc. Roy. Soc. Lond.*, 1922, **101A**, 94–100.

(To be concluded)

[MS. received November 19, 1927]

THE INFERENCE OF MIND

BY DONALD KEITH ADAMS

Institute of Psychology, Yale University

Few theoretical questions in psychology have occasioned as much controversy as that of the basis and method for the inference of mind in animals, and few have had as diverse answers. At one extreme are those for whom the animal mind is a whimsy as fantastic as any hobgoblin of fiction. At the other extreme are those who describe the mental processes of a creature so different from themselves as an ant, with an assurance that most psychologists would covet for the description of another human mind. Neither of these extremes, however, requires an exhaustive critique, and both will be ignored in the present inquiry. Further, all question of the interest, value or merit of the investigation of the animal mind to any individual or to society is here considered to be irrelevant. These being assumed, our problem is purely that of method.[1]

Among those who investigate the mental life of animals, it is commonly asserted that the fundamental assumptions of animal psychology are no more than extensions of the assumptions of psychology in general. One can observe directly the processes of only one mind, his own. All of us, however, conduct the practical affairs of life on the assumption

[1] Most of the general treatises of animal psychology discuss some of the questions here considered, but there are few journal articles dealing specifically with the subject of *safeguarding* the method of analogy, assuming its validity. Some relevant papers, not mentioned elsewhere in this article, are: Carr, H., 'The Interpretation of the Animal Mind,' PSYCHOL. REV., 1927, 34, pp. 87–106; Claparède, E., 'Les animaux, sont-ils conscients?' *Rev. phil.*, 1901, 51, pp. 481–498, and 'La psychologie comparée, est-elle légitime?' *Arch. de psychol.*, 1903, 5, pp. 13–35; Piéron, H., 'Les methodes de la psychologie zoologique,' *Rev. phil.*, 1904, 58, pp. 171–172; Read, Carveth, 'The Comparative Method in Psychology,' *Brit. J. Psychol.*, 1913, 6, pp. 44–59; Wheeler, O. A., 'The Basis of Comparative Psychology,' *Sociol. Rev.*, 1913, 6, pp. 338–347; Yerkes, R. M., 'Behaviorism and Genetic Psychology,' *J. Phil., Psychol. and Scient. Meth.*, 1917, 14, pp. 154–160.

that other humans, who look and act much as we ourselves look and act, have much the same sort of mental life as we enjoy. We rarely make this assumption explicit, nor need it, for our practical concerns, be more definite. The psychologist, however, is curious about the workings of other people's minds and for the purpose of satisfying his curiosity—more or less tentatively, for scientific curiosities are never fully satisfied—he makes this every day assumption somewhat more explicit and specific. He says, *To the degree that two or more people in similar situations behave similarly, their mental processes resemble each other.* It will be noted that the restriction to the class of humans implies that a certain degree of structural similarity is necessary to the predication of mental similarity. It should also be noted that the psychologist rarely troubles himself about his limiting cases, but leaves the definition of the class *humans* to the reader. That is to say, he does not specify in quantitative terms the degree of structural similarity that will permit the inference of mental similarity. Sometimes, indeed, he circumscribes his field more narrowly, but no more definitely, as when he writes a psychology of the normal, civilized, adult, human mind. It has happened upon occasion that some of his inferences as to other minds than his own have been limited to the minds of psychologists trained in the same laboratory. They are none the less valid for that reason, within those limits.

Some psychologists, noting the continuity of organismic structure, and curious about the minds of non-human animals, see no reason why the psychological assumption should not be extended to organisms in general, and proceed to work on that basis. The assumption upon which their inferences are based may be phrased as follows: *To the degree that two or more organisms resemble each other in structure, and in similar situations behave similarly, their mental processes resemble each other.* Now all three of the 'objective' similarities and resemblances are continuous variables: structure, situation, or behavior may be very unlike or almost indistinguishable from other structure, situation, or behavior, with any inter-

mediate degree of resemblance. The inference of mental processes in any other organism than oneself—including other humans—must always, of course, be a matter of greater or less probability, and it would seem perfectly futile to be concerned about the difficulty of selecting a phylogenetic point 'where mind begins.'

From the preceding considerations, which summarize the usual theoretical basis for the investigation of the animal mind, certain other circumstances become evident. It is plain that, their fundamental assumptions being essentially the same, the methods of investigation of the human and animal mind will be essentially the same. While we may not agree with Titchener that "We do, then, make the animals attend to stimuli and report their experiences to us; we do, after a fashion, make them introspect,"[2] we may with perfect propriety consider language as simply *a peculiarly exact and fluid way of behaving similarly*. We are frequently confronted by its inadequacies, but by comparison with any other mode of behavior, it is capable of the most exquisite exactitude of duplication. When the subject of a psychophysical experiment says 'heavier' he is behaving more or less exactly as the experimenter in a similar situation would behave if he were having a certain sort of experience, and the latter infers that the subject is having a similar experience. All our 'knowledge' of other minds is obtained through comparison—more or less conscious and formal—of other behavior with our own, and by inference from analogy. This may be called the general method of comparison, but to avoid a regrettably common confusion, is better known as the general method of analogy. This may be the psychologist's fallacy, but it is at the same time the sole avenue of escape from his solipsistic predicament.

So far, then, there is nothing unique about the method of investigation of the animal mind. Information about the minds of animals is obtained in the same way as information about the minds of other humans, namely, by analogy with one's own. At this point it is necessary to differentiate

[2] Titchener, E. B., 'A Text-book of Psychology,' New York, 1910, p. 33.

animal psychology, comparative psychology, and genetic psychology. The fact that the term comparative psychology is so commonly used synonymously with animal psychology indicates the prevalence of the confusion alluded to above, between the general method of analogy, by which all the data of all psychology are obtained, and the comparative method, which has to do with the treatment of the data of any science, no matter how obtained, whether by direct inspection or by inference. By comparative method is meant description in terms of likenesses and differences between or among individuals or classes of individuals; the classes, obviously, being products of prior comparison of individuals. Thus the 'comparative' prefixed to psychology means just what it means when prefixed, for example, to anatomy. Even Lloyd Morgan, in his careful chapter on 'Other minds than ours,' fails to keep clear the distinction between the analogical method of obtaining psychological data, and the comparative method of describing any data, with considerable impairment to his treatment.

In the application of the comparative method to its data, psychology is at no disadvantage as compared with any other science. But in the application of the general method of analogy to the *collection* of its data, animal psychology must contend with much greater difficulties than those besetting human psychology. For the nature of its subjects precludes the most facile and efficient form of behaving alike, namely, language.

Thus, by animal psychology is meant simply the psychology of non-human animals. Comparative psychology means psychological description in terms of likenesses and differences, and this may include the subject matter of both human and animal psychology. Genetic psychology means psychology from the point of view of genetic relationships in the psychic development either of individuals or of species. It should be noted that only the third of these *requires* the postulation of evolution.

To recapitulate and specify: we are interested in determining the mental contents and processes of animals, for

example, of cats. In other words, we wish to know what experiences, or more particularly, what *sorts* of experiences these animals have. Still more particularly, we wish to know whether or not they have certain complexly conditioned sorts of experiences, for example, ideation. These experiences cannot be directly observed, any more than one person can observe the experiences of another. Therefore we shall resort to the same expedient that we use in discovering the experiences of another person, except that we shall be forced to be more careful and critical because these creatures are probably more different from us than any human. We shall, then, observe the behavior of the animals in carefully controlled and fully described situations. We shall describe this behavior as fully and 'objectively' as may be.[3] Ideally, we then observe the behavior of humans in similar situations, and ask these to describe their experiences in detail, in language. Actually, this is rarely done. Usually the experimenter just imagines himself, his behavior, and his experiences in the situations he requires his animals to face, and draws his inferences by analogy from similarities between the animal's actual and his own imaginary behavior. In order to check his inferences the reader must do some imagining for himself.

Reduced to these bald terms, the method of inference by analogy from one's own mind to that of a non-human animal has seemed to many to be futile, not to say ridiculous. The extension of the general psychological assumption to animals appears to them the *reductio ad absurdum* of all psychology. Hence behaviorism. How can notions arrived at by such

[3] Some form of automatic graphic record is always desirable. Among those presently in use, the motion picture may be mentioned. The behavior of animal subjects in the experimental situations may be photographed from several points, and a complete and 'objective' record obtained. This technique may also be used for recording the reactions of humans to the same situations, so that anyone seeing the two records might make his own comparison, decide for himself whether or not there was sufficient similarity of behavior to warrant inference of similarity of experiences or mental processes, consult the protocols of the human subjects, and make such inferences as might be warranted. In the present state of our knowledge, it is highly desirable that some such technique as this should be employed, especially for the investigation of complex processes.

devious procedures, and so dependent upon the prejudices of the individual, be signified by inclusion among the data of *science?*

To others, the evident difficulties and superficial absurdity of this—the only—method of investigation of the animal mind have only made it more intriguing, and they have set about the probably impossible task of making the method foolproof, by the addition of certain restrictive provisions. (The most widely accepted of these is the celebrated principle or canon of Lloyd Morgan, which it is my purpose to examine.) On the basis of these a considerable amount of work has been done, and this cannot be disposed of by reading it summarily out of the field of science on the ground that it is not directly susceptible to all or even to any of the tests of practical application used in the physical sciences. The gravest charge that can be brought against this work is that of incomplete candor. For the investigation of the animal mind is usually justified on the ground that its method is essentially the same as that of psychology in general, arguing along the lines developed above. But as a rule, the fact, for example, that the inferences are made on the basis of similarities between the animal's recorded behavior and the *imagined* behavior of the experimenter is not explicit. This is not to be construed as meaning that this procedure is necessarily bad. In all probability it is often better than actually performing the experiments upon humans, as stipulated in the method; for it is usually necessary to abstract certain elements from the human experience anyway, and this can be done only in imagination. But this departure from the letter of the method should be noted explicitly.

The foregoing is only one of a whole nest of difficulties, some practical and some methodological, that infest the way of animal psychology. Obviously, the more of these that are explicitly recognized, the more circumspect will be the inferences made. Lloyd Morgan, the pioneer systematizer and methodologist of this field, early recognized some of them and sought in his celebrated canon to eliminate them. "In no case may we interpret an action as the outcome of the

exercise of a higher psychical faculty, if it can be interpreted as the outcome of the exercise of one which stands lower in the psychological scale." [4] This is plainly intended as an adaptation to the problems of animal psychology of the general Law of Parsimony, first formulated by William of Occam: "*Entiae non sunt multiplicanda praeter necessitatem.*" Now this law, or rather principle of scientific method, has been one of the most useful tools of inductive science in pruning its concepts, evaluating hypotheses and comparing theories. It simply expresses the drive of all science towards unity, the desire to reduce a complex phenomenal manifold to a single theoretical or conceptual unity.

Morgan's canon, however, instead of being as commonly considered, a special case of the law of parsimony, is not related to it, and may on occasion work to exactly opposite effect. It involves a considerable number of assumptions, of some of which Morgan was quite aware. "But what, it may be asked is the logical basis upon which this principle is founded? If it be true that the animal mind can be interpreted only in the light of our knowledge of human mind, why should we not use this method of interpretation freely, frankly, and fully? Is there not some contradiction in refusing to do so? For, first, it is contended that we must use the human mind as a key by which to read the brute mind, and then it is contended that this key must be applied with a difference. If we apply the key at all, should we not apply it without reservation?" He is here raising a point that present criticism is not likely to raise, but his answer to this requires the statement of some of his assumptions.

"This criticism might be valid if we were considering the question apart from evolution. Here evolution is postulated. The problem is this: (1) Given a number of divergently ascending grades of organisms, with divergently increasing complexity of organic structure and correlated activities: (2) granted that associated with the increasing organic complexity there is increasing mental or psychical complexity:

[4] Morgan, C. Lloyd, 'Introduction to Comparative Psychology,' London, 1903, 2d ed., p. 53.

(3) granted that in man the organic complexity, the complexity of correlated activities, and the associated mental or psychical complexity, has reached the maximum as yet attained: (4) to gauge the psychical level to which any organism has been evolved. As we have already seen, we are forced, as men, to gauge the psychical level of the animal in terms of the only mind of which we have first-hand knowledge, namely the human mind. [If he adhered to the logic by which he justifies animal psychology, he would here be compelled to say, not 'the human mind,' but 'my mind.'] But how are we to apply the gauge?" [5]

Morgan then launches into a discussion, rather abstract and irrelevant, of ways of applying this gauge—a discussion which is obscured rather than illuminated by his diagrams of the 'Method of Levels,' the 'Method of Uniform Reduction,' and the 'Method of Variation,' and which betrays several more assumptions.

One writer after quoting the passage quoted above, leaves the subject with this neat comment: "Thus we assume, on the evolutionary hypothesis, greater simplicity of mind in the lower animal forms, and then adopt the principle of interpretation which is capable of giving results consistent with that premise." [6] This consideration alone should be sufficient to discredit the principle. The point I wish to make is that, regardless of their probability, this string of assumptions and the others implicit in them, are *wholly unnecessary*. Here is Morgan trying to adapt the law of parsimony to psychology, and violating it in the same breath by 'multiplying entities,' making quantities of unnecessary assumptions. Among those explicit and implicit, in the above quotation and the subsequent discussion, there seem to be these: evolution is assumed to be orthogenetic; differences between organisms are assumed to be entirely differences of degree of complexity; the same assumption is made for their associated action systems and for their 'psychical levels.' But it is unnecessary to add to the list he gives. His fundamental postulate, that

[5] *Op. cit.*, p. 55.
[6] Article on 'Animal Psychology,' *New International Encyclopedia*, 1, p. 661.

of evolution, is no more necessary to the inference of the animal mind than it is to the inference of another human mind.

So much for the logical and methodological sins of Lloyd Morgan's canon, though the preceding list is far from exhausting them. Let us turn to some of its practical difficulties. In the first place, how are we to tell of two alternative mental processes (it should be noted in justice to Morgan that he means by faculty, psychological process) which is the higher and which the lower? The only clue we get from him is that in a paraphrase of his principle, he substitutes for 'lower in the psychological scale,' 'lower in the scale of psychological evolution and development.' But that is precisely what he is trying to determine! In other words, if he knows what he needs to know in order to apply his canon, he already has the answer to his question, and consequently no need of the canon.

In the second place, his use of the word 'outcome,' and the whole tenor of his discussion, suggest, if they do not imply, a causal relation between the 'exercise of a psychical faculty' and the action which is its 'outcome.' Thus in addition to its intrinsic disabilities, the canon carries with it all the tangled difficulties of the concepts of cause and causal necessity. Finally, the phraseology of the canon is much less definite than it might be. No principle that aims to do what Morgan's canon aims to do can be perfectly definite, but a much better approximation to precision may be obtained.

The foregoing considerations seem sufficient to establish the inutility of Lloyd Morgan's canon to criticise the inference of mind. If, on the basis of the fundamental assumption of everyday social life, and of psychology—the general method of analogy—we are to infer the experiences or mental processes of animals, we must have a principle of interpretation that is at once more parsimonious, more rigorously defined, more workable, and wholly void of the vitiating connotations of cause. The following principle or rule of procedure is suggested as fulfilling these conditions: *Any experience or mental*

process in another organism can be inferred from structure, situation, history, and behavior only when a similar experience or mental process is or has been invariably associated with similar structure, situation, history and behavior in oneself; and the probability of the inference will be proportional to the degree of the similarity.

To anticipate questions that might arise, it is probably desirable at this point to indicate some of the things that this principle does not do. In the first place it does not supply a criterion of consciousness, of mind, or of anything else. It does not indicate any point in the continuum of similarity where the probability of inferences based on that similarity is attenuated to one half or less. It is not quantitative in the sense that it measures similarities in the units of any scale, although it requires judgments of greater and less. Thus while I can say that the structure, history and behavior of a cat resemble mine more than do those of a fish, and while I can say that morphologically I have all the mammalian characteristics in common with the cat, and only a few more than the vertebrate characteristics in common with the fish, I cannot say that the cat is so many times as like me as the fish, nor that the cat is so many units like me and the fish so many less. We never, of course, get perfect similarity between the inferring human and any other animal. The nearest approach to that is the case of identical twins, where inferences are sometimes made with astonishing accuracy, even without very similar histories or situations.[7] At what degree of similarity inference ceases to have any validity, cannot be determined in the abstract, but will have to be determined for each case from the descriptions given.

The principle here adopted makes certain descriptive demands that may need elaboration. Obviously, if the probability of our inferences depends upon similarity of structure, situation, history and behavior, adequate descriptions of these in both subject and experimenter are essential. Since our data are derived in this round-about fashion, by analogy, rather than by simple inspection, these descriptions

[7] Muller, H. J., 'Mental Traits and Heredity,' *J. Hered.*, 1925, 16, pp. 433–448.

cannot be omitted. We consider first the description of structure.

Other things than structure being equal, inferences as to the mental processes of another human are more likely to be correct than are inferences to those of a chimpanzee. The chimpanzee mind can be inferred with more assurance than the cat's, and the cat's more confidently than the frog's. But not all of our comparisons will be between animals so widely divergent in structure, and in finer comparisons the relative significance of various morphological features becomes important. Yerkes [8] lists in order of increasing importance, general form of organism, the nervous system, and specialization within the nervous system. This recalls the physiological tradition that the cerebrum is more intimately associated with mental processes than any other organ, and reaffirms the importance of the search for anatomical and physiological correlates of mental processes, a search, incidentally, demanding an integration of scientific inquiry that can never be achieved so long as we work in a science instead of on a problem.

There are, of course, practical limits to the description of structure. The experimenter cannot describe his own nervous system, but must assume a knowledge of human anatomy in his reader. Again, he is unlikely to be able to give more of the anatomy of his animal subjects than is known of the species of which they are members, and the reader can be referred to this. But such structural specializations as a prehensile organ, a fovea, vibrissae, etc., should be noted particularly. Again, individual peculiarities, where they or their significance can be detected, must be described. Thus for cats in a Thorndike puzzle box, the size of the head is significant, and if this varies materially from one individual to another, should be noted.

The description of situation also offers some difficulties, some of which are eliminated by definition. By situation I mean the 'external environment,' as perceived *by the experi-*

[8] Yerkes, R. M., 'Animal Psychology and Criteria of the Psychic,' *J. Phil., Psychol. and Scient. Meth.*, 1905, 2, pp. 141–149.

menter. Inasmuch as he has arranged the situations he will see more than the animal. But by defining the situation as 'the environment as seen by the experimenter,' the question of widening perceptions is referred to the proper category, namely that of history. Other things including history being equal, the probability of the inference of mind will be proportional to the similarity existing between the situation of the animal and the (actual or imaginary) situation of the experimenter.

The most difficult of the categories which are required to be similar, the one most usually neglected, and one of the most important, is that of history. A naïve, but prevalent, error is that of inferring mental processes—or their absence—in animals by comparing their behavior in a human and human-made situation, *i.e.*, one of which the experimenter has had abundant experience, with the experimenter's actual or imagined behavior in that situation. Thus, a situation that is perfectly familiar to a human, as the button fastening of Thorndike or the bolt fastening of Hobhouse, may be entirely without the cat's experience, and a superficial observer would say that the human would turn the button at once while the cat makes 'random' movements. Now it is probably impossible to get a human competent to report on his experiences who has had no experience with such fastenings. And the experimenter in making inferences from his imagined behavior, must abstract from his experience of similar fastenings, or the animal's behavior is bound to seem frightfully stupid. So far as the investigation of animal behavior is experimental, just so far it consists exactly in putting the animal into *our*, *human*, kinds of situation. We can do no other, for if instead we put ourselves into *their* kind of situation, we find ourselves ridiculously inadequate, and inferences as to their mental processes and experiences become mere fantastic and more or less groundless speculations, as witness the nature fakers, and those who are so overwhelmed by the spectacle of the animals doing what they cannot begin to do that they postulate the most marvelously complex and flexible instincts to explain them. Incidentally, this indi-

cates another limitation of our method: so far as animals do things that we cannot do, the mental processes and experiences associated with those activities are forever inaccessible to the method of inference from analogy, as limited by the principle here proposed, and one speculation is just about as good as another.

There is another source of error, working in the contrary direction, that results from the ignoring of history. If an animal *has* had experience of a given situation, or a similar one—this is too rarely definitely known—and his mental processes are inferred on the supposition that he has *not* had such experience, the inferences so made are worthless, or at least so attenuated in probability as to be safely negligible.

To some it may appear that the category of history is used to supplement other categories where they are deficient. We have been forced to define the situation as 'the physical environment as perceived by the experimenter.' We are unable to describe the situation as seen by the animal, but we can in a measure correct this deficiency by an adequate description of the animal's history. Again it might be urged that a really complete description of structure would include those (presently hypothetical) conditions of minute neural anatomy that are associated with past experiences, and that a description of history is simply a makeshift substitute for this. Be that as it may, a description of history is the best that can presently be compassed. It alone would immeasurably enhance many contemporary researches.

The description and comparison of behavior have, perhaps, more pitfalls for the unwary than any of the other categories. It must in the first place be 'objective.' The only meaning I can give to the term 'objective' is 'independent of any *particular* observer.' In the second place, behavior should be described so far as possible without reference to the possible or probable mental processes of the behaving animal. The inference of mental processes must wait until the descriptions of all the categories are completed. The above two propositions must be considered separate and distinct, and neither has anything to do with anthropomorphic or purposive

description. A description of behavior can be both anthropomorphic and purposive, without transgressing the limits imposed by the propositions. Thus, a cat is in a puzzle box. Various members, and the total bulk of the animal, are in motion. Suppose these movements—all of them—could be measured and described in terms of millimeters, degrees, radians, velocities, dynes, ergs, and what not. The fact remains that the animal's behavior is not described until it is stated, or the notion is in some way conveyed, that the animal is *trying to get out.* A motion picture would convey this notion at once to anyone, but where this satisfactorily 'objective' method of recording is not used, and recourse must be had to verbal description, no satisfactory or adequate picture of the animal's activities can be conveyed if this notion is omitted and not assumed by the reader. The most extreme objectivist assumes, no matter how he phrase it, that his cat is going to try to get out of his puzzle box. If the cat does not try, perhaps he will not experiment further with it, but he must tell us that the animal did not try, if not in those words, in others. Man *is* the measure of all things, and his effort completely to avoid anthropomorphism in describing the behavior of other animals is just as successful as his effort to lift himself by his own bootstraps. When he succeeds in the latter attempt we may expect to see him succeed in the former. Of course, where possible this anthropomorphic form of description is to be avoided. My point is that it is sometimes inevitable in the interests of adequate description, and that anthropomorphic or purposive description does not necessarily involve any loss of objectivity or any inference of mental processes. Behavior, as well as mind, can be described as conative.

To put the matter still differently: the whole method of psychology is anthropomorphic, or more properly *auto*-morphic. The aim of the present inquiry is to render this automorphism more self-conscious and critical. So far as possible, then, we will leave our automorphism out of our description of behavior, and bring it in, frankly labeled inference, and criticised by the principle here proposed, after

all the comparisons demanded by the principle have been made. However, there will be times when the adequate description of behavior will require a sort of naïve anthropomorphism, and at such times, we will neither shirk the description, nor couch it in periphrastic and disingenuous language in order to convince ourselves and our readers that we are not anthropomorphic.

The category of behavior must not be understood, moreover, to include only the specific segment of behavior from which it is proposed to infer experience. The total action repertory must be considered, even if not explicitly described. Thus our inference of the mental process associated with an act will be governed in part by our knowledge that that act is or is not common to the entire species represented by our animal subject.

There remains the category of experience or mental process. All the other categories require two descriptions, one of the animal and one of the experimenter. The experimenter can describe only his own experience. This crucial description is usually omitted. In many instances it can be omitted without seriously impairing the inference, for the experimenter does not often actually put himself into the situation and report his experiences, but simply imagines himself and his behavior in the experimental situation and ascribes to the animal his imagined experiences. As noted above, inasmuch as he must abstract, for example, the verbal constituents from his experience in any case, this procedure is not to be unreservedly condemned. However, it must be admitted that he thereby forgoes what might be very useful clues, and at the same time weakens his inferences. Sometimes, of course, the behavior is so clear cut, that he can say to the reader: "I could not do this and that without having such and such experiences. Could you?" This critical comparison is rarely made explicitly, and it is often unnecessary to make it so, but for many inferences of the animal mind it would constitute a most devastating *reductio ad absurdum*.

The relative importance to the inference of mind of similarities in structure, situation, history, and behavior

cannot be generally defined. Usually each inference must be made and criticised by the principle on its own peculiar merits, although, as we shall see, it is sometimes possible to make type inferences. It might be urged, however, that similarity of structure, for example, should be weighted most. In the first place, it is the observed continuity of structure that first suggests the possibility of inferring mind. Again, in one's own thinking, structure is usually the first thing one looks for in evaluating an inference of mental process. Thus, though the gross behavior of an ant in a given situation might be very similar to mine in a similar one, I should infer its mental processes with much less confidence than those of a chimpanzee whose behavior resembled mine less. But a similar case might be made for almost any of the categories required, and it is my opinion, inasmuch as all are required, and the probability of an inference is proportional to the similarity in all the categories, that they may in general be regarded as of equal weight.

A question that would arise more frequently in use is that of the mode of application of the principle. Can the criticism of inferences enunciated in this principle be applied to inferences of *types* of experiences, or must each inference be made and criticised specially and separately? I believe that we can abstract the common circumstances from many actual or hypothetical structures, situations, histories, and acts, and infer experiences from the similarity of these common circumstances to those abstracted in like manner from our own structure, situations, history and acts. We can certainly do this for certain relatively simple and common experiences associated with relatively well-defined and stereotyped behavior. Thus few persons, lay or scientific, would hesitate to infer the general experience of pain in cats after seeing one or two of them in even one of the situations that are usually associated with pain in themselves. I have occasionally stepped upon a cat's foot, observed his sudden withdrawal, heard his cries, seen the immediately following behavior of retreat and licking the stimulated member, and noted that for a greater or less period of time afterward he did not come

within reach of me. These observations, though not particularly quantitative, are 'objective,' and I do not hesitate to infer the experience of pain on this basis, knowing what I do of the animal's structure and history, for example, that it has free nerve endings in the skin and that it is not under an anæsthetic. Nor do I doubt that if I cut off an ear the animal would experience pain, though I have never myself been in that situation. Neither, though I have no such member, do I believe that his tail could be amputated without pain. In general, unless some structural, historical, or behavioral datum negatives it, I should think that any stimulus associated with the sensation of pain in my experience would be associated with pain in a cat's. In reaching this conclusion, it is not necessary to apply the principle to each instance, for we can abstract from all the pain-producing situations and structures the common feature of stimulation of certain types of nerve endings. We can abstract from the behavior the common element of withdrawal from the stimulus. And we infer the common experience of pain. Of course, as we have abstracted one or only a few of the circumstances, thus augmenting the extension of our inference, we have diminished its intension. We can say less about it, perhaps only that the experience is pain. According as the inference is detailed, then it must be special and individual. For certain experiences, those readily classified, and associated with structures, situations, histories and behavior that are readily classified, such general inferences may well suffice. But the border-line cases, which are the ones most likely to be experimented upon, must be examined and criticized specifically.

There is another difficulty, more or less irrelevant to the method, but inescapable in animal psychology. It has been noted that the experimenter in inferring an animal's experiences from his own, must abstract, among other things, the verbal components from his experience. Unfortunately, the difficulty does not end here, or it would be that only for the experimenter. But he can report the inferred nonverbal experiences and mental processes only in words. As

Hobhouse says, "Our difficulty in representing, or at least in trying to set forth his presumed state of mind in words, is precisely that in clothing it in words we are doing what he cannot do." [9] In the face of this, the most facile criticism of the behaviorist is to exclaim, "What a fool this man must be to suppose that animals think like this"! The interested psychologist will, of course, try to envisage the described experience without verbal components. Only thus can any essayed description of the animal mind seem other than the veriest moonshine.

SUMMARY

It has been found on methodological and practical grounds that Lloyd Morgan's canon is insufficient to criticize the inference of mind, and an alternative principle has been suggested: any experience or mental process in another organism can be inferred from structure, situation, history and behavior only when a similar experience or mental process is or has been invariably associated with similar structure, situation, history and behavior in oneself; and the probability of the inference will be proportional to the degree of the similarity. This is believed to indicate more precisely the directions that are taken or that should be taken by criticism of inferences of mental process or experience. The various descriptions demanded by the principle, and the modes of its application, are explicated in the ensuing discussion.

[MS. received December 23, 1927]

[9] Hobhouse, L. D., 'Mind in Evolution,' 2d ed., 1915, p. 321.

A FURTHER EXPERIMENTAL STUDY OF THE DEVELOPMENT OF BEHAVIOR

BY LEONARD CARMICHAEL

Brown University

PART I. INTRODUCTION

The present writer has presented in two previous papers certain experimental findings upon the development of behavior in organisms artificially removed from the influence of external stimulation.[1] In this paper some additional results supplementary to those already reported will be given.

The general problem which the experiments treated here, as well as those already reported, was set to solve is that of the relative importance of heredity and environment in the development of behavior.

As has been suggested before, much of the earlier experimental work upon this problem has been at least partially invalidated by the fact that it was based on the assumption that all of the behavior displayed by an organism at the time of birth or hatching from an egg is native. This assumption is unsound because of the fact that before birth or hatching the organism in question may have been actively conditioned to stimuli of many sorts. In presenting the evidence upon this matter, the experimental work of Spalding,[2] Craig,[3]

[1] Carmichael, L., 'The Development of Behavior in Vertebrates Experimentally Removed from the Influence of External Stimulation,' PSYCHOL. REV., 1926, 33, pp. 51–58. Idem, 'A Further Study of the Development of Behavior in Vertebrates Experimentally Removed from the Influence of External Stimulation,' PSYCHOL. REV., 1927, 34, pp. 34–47. The experiments recorded here were carried on in the Psychological Laboratory of Princeton University.

[2] Spalding, D. A., 'Instinct, with Original Observations on Young Animals,' *Macmillan's Mag.*, 1873, 27, pp. 282–293.

[3] Craig, W., 'Observations on Doves Learning to Drink,' *J. Anim. Behav.*, 1912, 2, pp. 273–279.

Yerkes,[4] Breed,[5] Shepard,[6] Moseley,[7] Brown,[8] Bird,[9] Lane,[10] Stone,[11] Minkowski,[12] Coghill,[13] and others has been reviewed. On the basis of the evidence presented by these experimentalists, the conclusion is drawn that if a given act of behavior is to be shown definitely not to be acquired, it must appear the very first time that appropriate stimulation is applied to the neuromuscular mechanism involved. The technique of the experiments already reported has made it possible at least approximately to secure this condition.

In the first series of experiments, the embryos of the frog (*Rana sylvatica*) and of the salamander (*Amblystoma punctatum*) were used. At a stage much earlier in embryonic development than that at which any observable activity begins, the experimental embryos were placed in a solution of chloretone. Previous work had shown that living organisms placed in a solution containing certain concentrations of this drug continue morphological growth, but do not display any bodily movement. At the same time that the animals in the experimental groups were placed in the solution, similar control groups of embryos were placed in development dishes

[4] Yerkes, R. M., 'The Heredity of Savageness and Wildness in Rats,' *J. Anim. Behav.*, 1913, 3, pp. 286–296.

[5] Breed, F. S., 'The Development of Certain Instincts and Habits in Chicks,' *Behav. Monog.*, 1911, 1, pp. 2*ff*.

[6] Shepard, J. F., and Breed, F. S., 'Maturation and Use in the Development of an Instinct,' *J. Anim. Behav.*, 1913, 3, pp. 283*ff*.

[7] Moseley, D., 'The Accuracy of the Pecking Response in Chicks,' *J. Comp. Psychol.*, 1925, 5, pp. 75–97.

[8] Brown, T. G., 'On the Activities of the Central Nervous System of the Unborn Fetus of the Cat, with a Discussion of the Question whether Progression (Walking, etc.) is a "Learnt" Complex,' *J. Physiol.*, 1915, 49, pp. 208–215.

[9] Bird, C., 'The Relative Importance of Maturation and Habit in the Development of an Instinct,' *Ped. Sem.*, 1925, 32, p. 81.

[10] Lane, H. H., 'The Correlation Between Structure and Function in Development of the Special Senses of the White Rat,' A Dissertation, Norman, Oklahoma, 1917.

[11] Stone, C. P., on congenital behavior exhibited by prenatally delivered guinea-pigs, at present unpublished.

[12] Minkowski, M., 'Zum gegenwärtigen Stand der Lehre von den Reflexen,' Neurologische und psychiatrische Abhandlungen aus dem *Sweizer Arch. f. Neur. u. Psychiat.*, 1925, Heft I.

[13] Coghill, G. E., 'Correlated Anatomical and Physiological Studies of the Growth of the Nervous System of Amphibia,' *J. Comp. Neur.*, 1914, 24; *Ibid.*, 1916, 26; *Ibid.*, 1926, 40.

containing tap water. After the embryos of these second control groups had so far developed that they were free swimming organisms, the animals in the experimental groups still remained quite immobile. These drugged organisms showed approximately normal anatomical development but they did not respond even to vigorous external stimulation. On placing the anæsthetized embryos in fresh water, however, they soon showed signs of movement when stimulated, and in a few minutes they were swimming in a manner not to be distinguished from that of the individuals of the control group which had been allowed to develop normally.

These experimental results, nevertheless, did not present final evidence upon the question of the nativistic determination of the structures subtending behavior. A certain time always elapsed after the organism was placed in fresh water before it showed any movement in response to continued external stimulation. This time might, therefore, be considered either as an interval of very rapid learning or as a period necessary for the removal of the masking effect of the drug. A second series of experiments was therefore carried out in an effort to study the time required for the elimination of the effect of the anæsthetic. In these studies, the experimental organisms were treated as in the first series just reviewed, and the time required for the first release of each embryo from the anæsthetic in which it had developed was definitely recorded. These animals were then allowed to swim about freely in separate small dishes for a day and they were then reanæsthetized and kept under the influence of chloretone for twenty-four hours. These same individuals were then released from the drug for the second time. The times required for this process were recorded for each animal. It was found on comparing this time necessary for the second release with that required for the initial release that there was no significant difference. From this result, it was concluded that the time elapsing between the first removal of the organism from the anæsthetic in which it had been kept and the appearance of the very first movement in response to stimulation was probably to be regarded as a period necessary

for the elimination of the effect of the drug and not as a time of rapid learning.

In a theoretical consideration of these results, it was pointed out, however, that the data presented by no means requires the acceptance of the crude nativistic maturation hypothesis. This is made sure by the fact that all that is known concerning the actual developmental mechanics of the nervous system makes it seem that if the stimulation and response of the nervous elements themselves is eliminated, there is no neural development. Moreover, it seems probable that this internal stimulation and response of the nervous system must, in the first place at least, be initiated by environmental stimulation. It seems, however, that once initiated, this process of development may continue for a time at least in relative independence of external stimulation.

Part II. Experimental

The essentially abnormal situation conditioning the development of an organism raised under an anæsthetic made it seem desirable to attempt a check upon the results secured in this manner by some other technique. If external stimulation is as important in the development of the response mechanism as many writers have claimed,[14] it seems that there should be retardation or abnormality in the development of organisms raised under conditions that are relatively free from the usual stimulus changes of the normal environment.[15]

In the Princeton Psychological Laboratory, there is an excellent sound- and light-proof room. This chamber may be so isolated that it is absolutely dark and almost completely vibration-proof.[16]

[14] *Cf.*, for example, Burnham, W. H., 'The Significance of Stimulation in the Development of the Nervous System,' *Amer. J. Psychol.*, 1917, **28**, pp. 36–56; and Carmichael, L., 'Heredity and Environment: Are They Antithetical?,' *J. Abn. & Soc. Psychol.*, 1925, **20**, pp. 245–260. Both of these papers present relevant bibliography.

[15] Wells, W. R., 'The Meaning of "Inherited" and "Acquired" in Reference to Instinct,' *J. Abn. & Soc. Psychol.*, 1922, **17**, pp. 159*ff*.

[16] Langfeld, H. S., 'Princeton Psychological Laboratory,' J. Exper. Psychol., 1926, **9**, pp. 259–270.

As a check upon the work carried out on organisms developed under an anæsthetic, certain embryos of the salamander (*Amblystoma punctatum*) were allowed to grow in the isolation of this room. Those developed in this artificial situation were then compared with organisms allowed to grow in a relatively normal environment.

The Amblystoma eggs were prepared as in previous experiments. They were then divided into three sets and placed in development dishes. The dishes containing the first of these sets, which will here be called the experimental group, were placed on a deep pad of cotton batting on the cork floor of the sound- and light-proof room. The four doors of this room were then closed, and the outer door sealed. One of the other sets of organisms was placed in a little used part of the laboratory, but in a room which was subject to the ordinary diurnal changes of light, temperature, and incidental vibration. The third group was placed in the laboratory shop where the organisms which composed it were subject not only to changes in temperature and light, but to the vibrations incident to the operation of the shop machines and tools.

No difference could be noted between the development of the organisms in these two latter groups, and when the individuals in both these sets were freely swimming, the outer door of the isolation chamber was unsealed. The experimenter, after closing the outer doors so as to keep the inner room still completely dark, entered the isolation chamber with as little vibration as possible, holding an unlighted electric 'flashlight' in his hand. When he was directly over the development dishes, the light was switched on and several of the salamanders darted off with the same rapid swimming movements which characterize these organisms when normally developed.

It is certain that this small electric light was the first photic stimulus which had affected these organisms since a very early egg-cleavage stage. These embryos had, moreover, been isolated as well from almost all other vibrations. In spite of the difference in the stimuli affecting the three

groups of organisms during development, no difference in behavior could be noted between the individuals of the experimental group and those of the two check groups which developed in the outer laboratory.

The evidence of these experiments seems to substantiate the results of our previous work done upon the development of behavior. It seems that immediate external stimulation is not necessary for the development of those neuromuscular mechanisms which subtend behavior in the types studied.

Since the publication of the previous papers on this subject, an additional experimental group of embryos has also been developed under chloretone. The individuals in this group were released one by one and an effort was made to accurately observe the very earliest movements which took place as the organism gradually emerged from the anæsthetic. The early movements of the normal salamander embryo have been already described by Coghill.[17]

The observation of the earliest movements of the organisms as they gradually 'come out of' the anæsthetic in which they have developed is not easy. The technique consisted in removing a drugged organism from the chloretone solution by means of a pipette and placing it in a clean development dish containing fresh water. The experimenter, while observing the animal under a lens, gently touched the various points on the bodily surface of the organism with a small blunt glass rod. This stimulation continued until the first responses were noted. Some twenty-four organisms were studied in this manner. The protocols collected during this study are difficult to interpret in regard to the exact nature of first response to external stimulation made by these organisms as the anæsthetic is eliminated. In general, however, the first movement seems to have been a twitching or slight movement of the head or tail. These first twitches sometimes seemed to be toward the side stimulated and sometimes away from it. So far as the experimenter could

[17] Coghill, G. E., 'The Development of the Swimming Movement in Amphibian Embryos,' *Anat. Rec.*, 1908, 2, p. 148; and Herrick, C. J., and Coghill, G. E., 'The Development of Reflex Mechanisms in *Amblystoma*,' *J. Comp. Neur.*, 1915, 25, esp. p. 72.

determine, the orientation of the response in regard to the point of stimulation could not be accounted for in terms of the vigor of stimulation. Sometimes the head or tail apparently indiscriminately would move toward or away from a comparatively light or heavy touch. It seems, therefore, that it may be concluded on the basis of these tentative observations, that the first movements after release of these organisms were 'random.' The word 'random' is here used in the sense of undetermined by the place and nature of the stimulation. The response that is actually elicited must in any given trial be determined by the nature of the neuromuscular mechanism.

In general, however, the movements which followed these first twitches were similar to those already described by Coghill in the papers cited as characteristic of the first movements of the normally developed Amblystoma.

Part III. Conclusion and Summary

Nothing in the experiments reported in this study is to be taken as contradicting the theoretical conclusions suggested in earlier papers. It seems even more certain than it was before that a neuromuscular apparatus that has never before functioned as a response mechanism is capable of determining external behavior the very first time that it is effectively stimulated. It is difficult to see how the facts recorded here can be explained except on the assumption that heredity and environment are *interdependently* involved in the perfection of behavior. It is true, however, that these experiments point to the fact that there may be great variations in the necessary amount of environmental stimulation required to develop typical behavior patterns.

In conclusion, therefore, it may be said that in the present experiment salamanders were raised from an early and embryonic stage in such a manner as to strictly limit the possibility of environmental stimulation and response. This limitation was accomplished by allowing the organisms to develop either under an anæsthetic or to grow in a specially adapted isolation chamber. The results show that the de-

velopment of behavior occurs without constant and continuous response to external stimulation. A study was also made of the first movements shown by the released embryos. These movements were found to be 'random' in regard to the point of external stimulation, and in general to be not unlike those reported by Coghill as characteristic of the normally developed embryo. A knowledge of the developmental mechanics of the nervous system, moreover, precludes the use of these results as a final confirmation of the theory that the development of behavior is merely a maturation of native factors. On the contrary, viewed in its largest aspects, the development of behavior seems to be the result of the *interdependent* influence of the action of both environmental and hereditary factors.

[MS. received November 26, 1927]

Psychological Review Publications

Original contributions and discussions intended for the Psychological Review should be addressed to

Professor Howard C. Warren, Editor Psychological Review,
Princeton University, Princeton, N. J.

Original contributions and discussions intended for the Journal of Experimental Psychology should be addressed to

Professor Madison Bentley, Editor Journal of Experimental Psychology,
University of Illinois, Urbana, Illinois.

Contributions intended for the Psychological Monographs should be addressed to

Dr. Raymond Dodge, Editor Psychological Monographs,
Kent Hall, Yale University, New Haven, Conn.

Reviews of books and articles intended for the Psychological Bulletin, announcements and notes of current interest, and *books offered for review* should be sent to

Professor S. W. Fernberger, Editor Psychological Bulletin,
University of Pennsylvania, Philadelphia.

Titles and reprints intended for the Psychological Index should be sent to

Professor Walter S. Hunter, Editor Psychological Index,
Clark University, Worcester, Mass.

All business communications should be addressed to

Psychological Review Company, Princeton, N. J.

DIRECTORY OF
American Psychological Periodicals

American Journal of Psychology—Ithaca, N. Y.; Cornell University. Subscription $6.50. 624 pages ann. Ed. by M. F. Washburn, K. M. Dallenbach, Madison Bentley and E. G. Boring. Quarterly. General and experimental psychology. Founded 1887.

The Pedagogical Seminary and Journal of Genetic Psychology—Worcester, Mass.; Clark University Press. Subscription $7.00. 700 pages ann. Ed. by Carl Murchison. Quarterly. Child behavior, differential and genetic psychology. Founded 1891.

Psychological Review—Princeton, N. J.; Psychological Review Company. Subscription $5.00. 480 pages annually. Bi-monthly. General. Founded 1894. Edited by Howard C. Warren.

Psychological Bulletin—Princeton, N. J.; Psychological Review Company. Subscription $5.50. 720 pages annually. Psychological literature. Monthly. Founded 1904. Edited by Samuel W. Fernberger.

Psychological Monographs—Princeton, N. J.; Psychological Review Company. Subscription $6.00 per vol. 500 pp. Founded 1895. Ed. by Shepherd I. Franz. Published without fixed dates, each issue one or more researches.

Psychological Index—Princeton, N. J.; Psychological Review Company. Subscription $2.00. 300–400 pp. Founded 1895. Edited by W. S. Hunter. An annual bibliography of psychological literature.

Journal of Philosophy—New York; 515 W. 116th Street. Subscription $4. 728 pages per volume. Founded 1904. Bi-weekly. Edited by F. J. E. Woodbridge, Wendell T. Bush and H. W. Schneider

Archives of Psychology—Columbia University P. O., New York City. Subscription $6. 500 pp. per vol. Founded 1906. Ed. by R. S. Woodworth. Published without fixed dates, each number a single experimental study.

Journal of Abnormal Psychology and Social Psychology—Albany, N. Y. Subscription $5. Boyd Printing Co. Ed. by Morton Prince. In cooperation with Henry T. Moore. Quarterly. 432 pages ann. Founded 1906. Abnormal and social.

Psychological Clinic—Philadelphia; Psychological Clinic Press. Subscription $2.50. 288 pages. Ed. by Lightner Witmer. Founded 1907. Without fixed dates (9 numbers). Orthogenics, psychology, hygiene.

Training School Bulletin—Vineland, N. J.; The Training School. Subscription $1. 160 pages ann. Ed. by E. R. Johnstone. Founded 1904. Monthly (10 numbers). Psychology and training of defectives.

Comparative Psychology Monographs—Baltimore; Williams & Wilkins Co. Subscription $5. 500 pages per volume. Edited by W. S. Hunter. Published without fixed dates, each number a single research.

Psychoanalytic Review—Washington, D. C.; 3617 10th St., N. W. Subscription $6. 500 pages annually. Psychoanalysis. Quarterly. Founded 1913. Ed. by W. A. White and S. E. Jelliffe.

Journal of Experimental Psychology—Princeton, N. J. Psychological Review Company. 480 pages annually. Experimental. Subscription $5.00. Founded 1916. Bi-monthly. Ed. by Madison Bentley.

Journal of Applied Psychology—Baltimore, Md.; Williams & Wilkins Co. Subscription $4. 400 pages annually. Founded 1917. Quarterly. Edited by James P. Porter and William F. Book.

Journal of Comparative Psychology—Baltimore; Williams & Wilkins Company. Subscription $5. 500 pages annually. Founded 1921. Bi-monthly. Edited by Knight Dunlap and Robert M. Yerkes.

Genetic Psychology Monographs—Worcester, Mass.; Clark University Press. Subscription $7.00 per volume. Two volumes per year, 600 pages each. Ed. by Carl Murchison. Monthly. Each number one complete research. Child behavior, differential and genetic psychology. Founded 1925.

Psychological Abstracts—Eno Hall, Princeton, N. J. Edited by W. S. Hunter. Subscription $6.00. Monthly. 700 pages annually. Founded 1927.

Journal of General Psychology—Worcester, Mass. Clark University Press. Subscription $7.00. 600–700 pages annually. Edited by Carl Murchison. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1927.